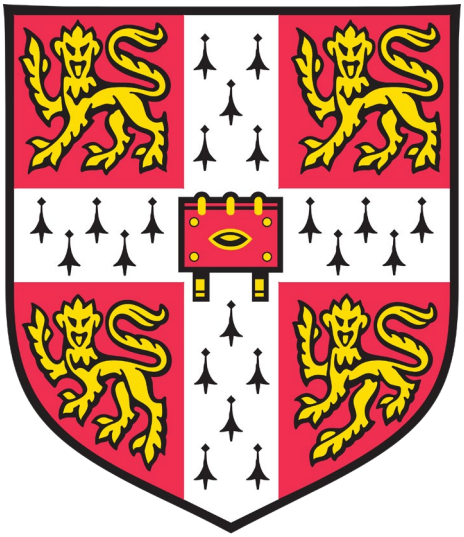
Investigating the Role of Stem-Loop 1 in the Assembly Process of HIV-1



Christopher James Hellmund

Jesus College

December 2017

This dissertation is submitted for the degree of Doctor of Philosophy

Summary

Name: Christopher James Hellmund

Thesis title: Investigating the Role of Stem-Loop 1 in the Assembly Process of HIV-1

Summary: An important step in the production of infectious HIV-1 particles is maturation of the virus core. This process is completed by cleavage of the capsid (CA) domain of Gag, from its precursor, CA-SP1, by the viral protease. Large deletions in stem-loop 1 (SL1) in the 5' untranslated region (UTR) of HIV-1 genomic RNA (gRNA) delay CA-SP1 processing. SL1 harbours the dimerisation initiation site (DIS) palindrome suggesting that efficient Gag processing may be linked to gRNA dimerization as shown in HIV-2. However, a dimerisation mutant with normal Gag processing was identified. Gag processing defects are hallmarks of late domain mutants, and SL1 mutation was found to result in reduced virus release. HIV-1 hijacks the host's endosomal complexes required for transport (ESCRT) pathway to enable budding. An ESCRT-associated protein, ALIX, is known to be capable of binding to the nucleocapsid (NC) domain of Gag using lipids or RNA as a 'bridge' in vitro. It was hypothesised that SL1 mutation disrupts an RNA-dependent interaction that occurs during virus assembly. Consistent with this, an intact SL1 was found to be required for efficient ALIX function. Increasing the abundance of gRNA in the cell by expressing it in trans accelerated CA-SP1 processing in a manner that required ALIX's binding motif in p6. Gag processing could also be accelerated by introducing previously identified compensatory mutations into the SP1 and NC domains of Gag, in a manner reminiscent of the actions of maturation inhibitor resistance mutations. The effects of the compensatory mutations were also dependent on intact late domain motifs. These data suggest that gRNA is involved in regulating virus budding and maturation through interaction with ALIX. A model is proposed whereby the packaging signal (psi) region of gRNA acts as a bridge between Gag and ALIX, acting as a checkpoint mechanism to promote Gag processing and optimise release of virions that have successfully packaged gRNA.

Declaration

This dissertation is the result of my own work and includes nothing that is the outcome of work done in collaboration. It is not substantially the same as any that I have submitted, or is being concurrently submitted, for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution. I further state that no substantial part of my dissertation has already been submitted, or is being concurrently submitted, for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution. It does not exceed the prescribed word limit for the relevant Degree Committee.

Acknowledgements

Firstly, I'd like to thank my supervisor Andrew Lever for giving me the opportunity to undertake this project, for supporting me to steer it in my own direction, and for giving me plenty of opportunities to present my work. My second supervisor, Bo Meng, has been a constant source of support and mentorship, always keeping me on my toes by challenging and discussing my scientific ideas as well as looking out for my general wellbeing. I've really enjoyed our philosophical chats about life over cups of tea and pints of beer.

All members of the Lever lab, past and present, have made it a fun place to work and I wholly blame them for turning me from a couch potato into a mountain climber and half marathon runner! Thanks to Eirini Vamva and Nick Norton for the PhD student solidarity, and to Nyarie Sithole for the daily words of wisdom. I am extremely grateful for the friendship and support I have received from Axel Fun, Mikaila Bandara, Carin Ingemarsdotter, Pratigya Gautam, Jack Hirst, Sophia Ho, Harriet Groom, HP Mok, Jing Garland, Julia Kenyon, and Benjamin Krishna from the Sinclair lab.

I'd like to thank my parents for supporting my unorthodox career choices so far, and for putting up with me as I returned to the nest to write this thesis. I'm also very appreciative to Amy and Shane who made me an excellent PhD survival kit which I think should be mass-produced and distributed to all PhD students.

I am extremely grateful for my friends back home in Bedford who have provided a much-needed link to the outside world, and to Katie Davies for the many hours of Gogglebox viewing accompanied by enormous quantities of food whilst we recovered from exhausting days in the lab.

And finally, although we only met when I was just starting to write my thesis, Louise Ainsworth has been an inspiration and has kept me motivated even on the toughest days.

Contents

Abbreviations

Abbreviation	Definition
AIDS	Acquired immune deficiency syndrome
ALIX	ALG-2 interacting protein X
ARS	Acute retroviral syndrome
ARV	Antiretroviral
BSA	Bovine serum albumin
CA	Capsid
CA_{CTD}	Carboxy-terminal domain of CA
CCLR	Cell culture lysis reagent
CCR5	C-C chemokine receptor 5
CD4	Cluster of differentiation 4
CHMP	Charged multivesicular body protein
CLIP	Crosslinking immunoprecipitation
CLIP-seq	Crosslinking immunoprecipitation-sequencing
CMV	Cytomegalovirus
CPSF6	Cleavage and polyadenylation specificity factor-6
CXCR4	C-X-C chemokine receptor 4
CypA	Cyclophilin A
DIG	Digoxigenin
DIS	Dimerisation initiation site
DLS	Dimer linkage structure
DMEM	Dulbecco's modified Eagle medium
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotide triphosphate
E. coli	*Escherichia coli*
ELISA	Enzyme-linked immunosorbent assay
Env	Envelope
FBS	Fetal bovine serum
FIV	Feline immunodeficiency virus
Gag	Group-specific antigen
gRNA	Genomic RNA
HAART	Highly active antiretroviral therapy
HIV (-1/-2)	Human immunodeficiency virus (type 1/type 2)
HRP	Horseradish peroxidase
HTLV	Human T-cell lymphotropic virus
IFN	Interferon
IN	Integrase
kb	Kilobase
kbp	Kilobase pairs
kDa	Kilodalton
LB medium	Luria Bertani medium
LDLR	Low-density lipoprotein receptor
LTR	Long terminal repeat
MA	Matrix
MBq	Megabecquerel
MDM	Monocyte-derived macrophage
MI	Maturation inhibitor
MLV	Murine leukemia virus

Abbreviation	Definition
MoMuLV	Moloney murine leukemia virus
M-PMV	Mason-Pfizer monkey virus
MVB	Multivesicular body
NC	Nucleocapsid
Nef	Negative factor
NLS	Nuclear localisation signal
NMR	Nuclear magnetic resonance
NPC	Nuclear pore complex
ORF	Open reading frame
PBS	Phosphate-buffered saline
Pbs	Primer binding site
PBMC	Peripheral blood mononuclear cell
PCR	Polymerase chain reaction
PIC	Preintegration complex
Pol	Polymerase
Psi/Ψ	Packaging signal
R	Repeat region
Rev	Regulator of expression of virion proteins
RNA	Ribonucleic acid
RRE	Rev response element
rRNA	Ribosomal ribonucleic acid
RSV	Rous sarcoma virus
RT	Reverse transcriptase
SDS	Sodium dodecyl sulfate
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
SHAPE	Selective 2'-hydroxyl acylation analysed by primer extension
siRNA	Short interfering ribonucleic acid
SIV	Simian immunodeficiency virus
SP1	Spacer peptide 1
SP2	Spacer peptide 2
ssRNA	Single-stranded ribonucleic acid
TAR	*Trans*-activation response element
Tat	*Trans*-activator of transcription
TBE	Tris-borate-EDTA
TBS	Tris-buffered saline
TIRFM	Total internal reflection fluorescence microscopy
tRNA	Transfer ribonucleic acid
TSAP	Thermosensitive alkaline phosphatase
TSG101	Tumour susceptibility gene 101
UTR	Untranslated region
U3	Unique 3' region
U5	Unique 5' region
Vif	Virion infectivity factor
VLP	Virus-like particle
Vpr	Viral protein R
Vpu	Viral protein U
vRNA	Viral RNA
VSV-G	Vesicular stomatitis virus G glycoprotein
XL-SHAPE	Crosslinking-SHAPE

List of figures

List of tables

1. Introduction

1.1. Overview

Human immunodeficiency virus type 1 (HIV-1) is a human retrovirus of the lentivirus genus and the causative agent of acquired immune deficiency syndrome (AIDS) (Barre-Sinoussi et al., 1983; Gallo et al., 1984; Popovic et al., 1984). AIDS can also less frequently be caused by HIV-2, although the viruses share little sequence homology (Guyader et al., 1987).

HIV-1 and HIV-2 are derived from distantly related simian immunodeficiency virus (SIV) strains which infect chimpanzees (SIV_{cpz}) (Huet et al., 1990) and sooty mangabeys (SIV_{smm}) (Hirsch et al., 1989) respectively. SIVs are generally asymptomatic in their natural hosts, indicating that the viruses and their hosts have co-evolved for a long period, a finding supported by phylogenetic analysis (Worobey et al., 2010). SIVcpz has however been shown to be associated with premature death in wild chimpanzees (Keele et al., 2009).

HIV-1 is a recent introduction into the human population. Molecular clock analysis estimates that the most recent common ancestor of all HIV-1 group M subtypes emerged between 1910 and 1930 in Kinshasa in the present-day Democratic Republic of Congo. Indeed Kinshasa is where the earliest known HIV-1 sequences were found, and today the city has the highest diversity of HIV-1 strains in the world (Faria et al., 2014; Korber et al., 2000).

HIV-1 quickly spread through high-risk groups in the population to create a pandemic that has infected an estimated 78 million people, of whom 35 million have died and 36.7 million are currently infected, with 2.1 million new infections occurring per year (UNAIDS, 2016). In the early days of the pandemic HIV infection was a death sentence due to lack of treatments, and patients died within years of becoming infected. Thanks to significant investment in research into antiretrovirals (ARVs) and the used of highly active antiretroviral therapy (HAART), in which combinations of ARVs are used to increase the barriers to resistance, people infected with the virus are now able to live healthy and long lives. Unfortunately, they must take these drugs for the rest of their lives to keep the virus under control, as cessation of treatment causes a rebound in viral load.

Efforts are intensifying to find a cure, although at present this still seems some way off. Furthermore, out of the 36.7 million people currently infected, only 18.2 million are accessing antiretroviral therapy (UNAIDS, 2016). Development of an effective vaccine is difficult because the most conserved epitopes on the virus's envelope (Env) glycoprotein are shielded by glycosylation and are embedded within a structure which undergoes conformational changes during binding to host cells (Barouch, 2009).

There are four lineages (termed "groups") of HIV-1 that arose from separate cross-species transmission events in West central Africa, likely through hunting and the bushmeat trade. HIV-1 group M is by far

the most successful, with a global geographical spread. Population bottlenecks have led to the establishment of nine subtypes, labelled A-D, F-H, J and K. Group M is responsible for most infections and deaths, whilst group O is responsible for less than 1% of infections, and groups N and P account for just a handful of documented cases restricted to Cameroon and the surrounding countries. Group M and N viruses are closely related to SIV found in chimpanzees, whilst group P and O viruses are most closely related to SIV found in gorillas (D'arc et al., 2015; Sharp and Hahn, 2011).

HIV is spread predominantly through heterosexual sexual intercourse, which accounts for 70% of transmission events. The remainder occurs through sexual intercourse between men who have sex with men, mother to child transmission (during pregnancy or at birth), intravenous drug use or, rarely now, blood and blood product transfusion (Shaw and Hunter, 2012).

Socioeconomic factors mean that the challenges posed by HIV-1 vary between world regions. In the developed world – where ARVs are more readily available – the focus is on preventing transmission by ensuring that infected individuals adhere to treatment, thus reducing the risk of transmission to their partners. An additional measure is pre-exposure prophylaxis, whereby ARVs are prescribed to individuals who are uninfected but are at high-risk of being infected, to prevent the establishment of infection if they are exposed to HIV. Post-exposure prophylaxis is administered following high-risk activity, and ARVs are also given to babies born to HIV-positive mothers. In developing countries, lack of education, access to diagnostic facilities and the availability of drugs are the major barriers to controlling transmission.

A cure for HIV infection remains elusive because although ARVs suppress the viral load, they cannot eliminate the reservoir of viruses that establish latency. Research is intensifying into approaches which will force the virus out of latency to enable it to become a target for pharmacological agents or immunotherapy – the so-called 'kick and kill' strategy (Archin and Margolis, 2014).

1.2. Clinical presentation of HIV infection

Clinically, HIV infection can be categorised into three stages (illustrated in Figure 1): the acute stage; the latent stage – not to be confused with viral latency; and AIDS. Acute HIV infection occurs during the peak of viremia in the first few weeks post-exposure, and is associated with a self-limiting flu-like set of symptoms termed acute retroviral syndrome (ARS) (Henn et al., 2017). These include fever, headache, malaise, cough and lymphadenopathy (Robb et al., 2016), meaning that ARS is symptomatically indistinguishable from a range of other infections.

The latent stage of infection is associated with few (if any) symptoms, and low levels of virus replication due to a robust immune response. In the absence of effective treatment with ARVs, replication in and killing of CD4+ T-helper cells results in a progressive decline in CD4+ cell levels, and consequently a decline in the ability of the immune system to control virus replication (Alizon and Magnus, 2012). This

stage of infection can last several years, but once CD4+ levels are sufficiently low the clinical signs of AIDS become apparent.

A weakened immune system enables establishment of infection and presentation of symptoms associated with opportunistic pathogens or malignancies that are normally cleared or controlled by healthy individuals. Opportunistic bacterial, fungal and viral pathogens associated with AIDS include *Mycobacterium tuberculosis*, *Streptococcus pneumoniae*, *Cryptococcus neoformans*, *Candida sp*., *Toxoplasma gondii*, and cytomegalovirus (CMV) (Holmes et al., 2003). Cancers commonly associated with AIDS include Kaposi's sarcoma and non-Hodgkin's lymphoma (Cheung, 2005). The inability of the immune system to combat these diseases results in rapid deterioration of health and short life expectancy for patients with AIDS (Morgan et al., 2002).

Figure 1 – Viral load and CD4+ T-cell count over the course of an untreated infection

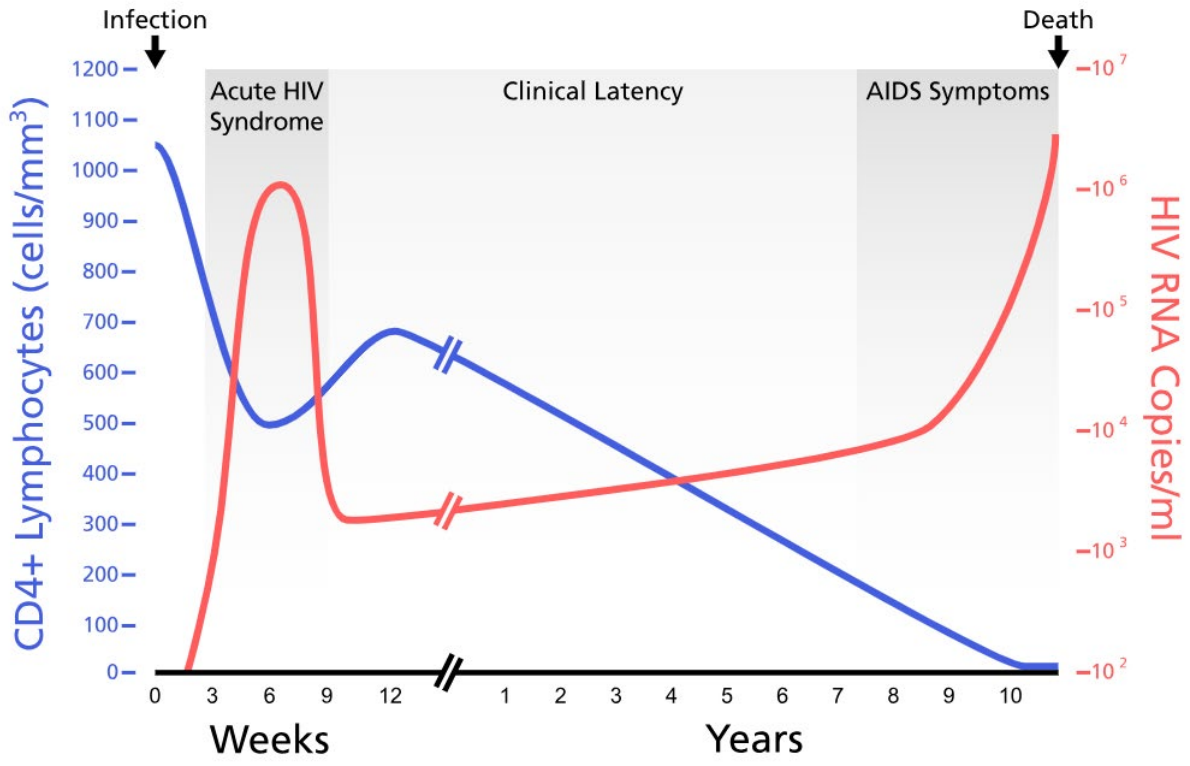


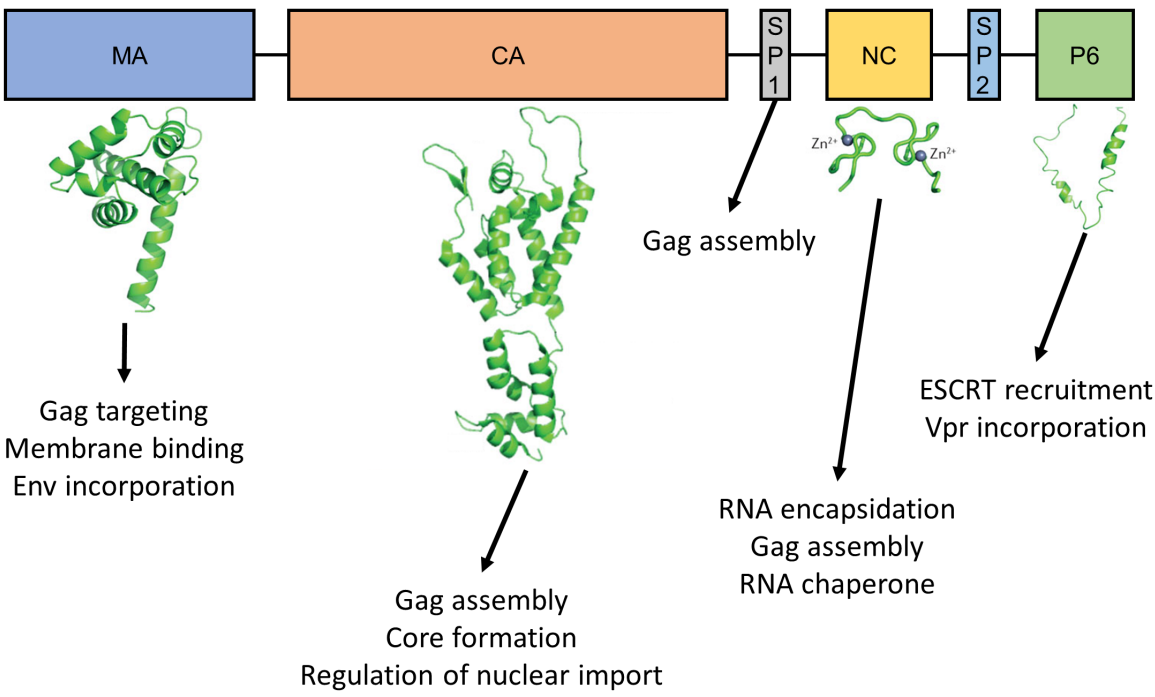
Illustration of the changes in CD4+ T-cell counts and viral load (measured by HIV RNA copies/ml) throughout the course of an infection in the absence of treatment. The infection can be divided into three stages. The initial acute phase lasts a matter of weeks and is associated with a rapid burst of viral replication accompanied by a decline in the CD4+ count, while patients may experience flu-like symptoms. A robust immune response brings the viral load under control and ushers in the latent stage of infection, which is generally asymptomatic. Over several years there is a gradual decline in the CD4+ cell count, which leads to AIDS, in which opportunistic pathogens and neoplasms cause a rapid decline in health, leading to death.

Source: https://commons.wikimedia.org/wiki/File:HIV-timecourse_simple.svg

1.3. HIV-1 genome organisation and structure

HIV-1 encodes 9 genes producing 3 structural, 2 regulatory and 4 accessory gene products. The structural proteins are group-specific antigen (Gag), polymerase (Pol) and Env. Gag and GagPol are multi-domain proteins that are processed into mature proteins in the late stages of viral assembly by the viral protease (PR). The domains of Gag and their functions in virus replication are shown in Figure 2.

Figure 2 – Structure and function of Gag domains



Box diagram illustrating domains of Gag (to scale), with crystal structures and the key functions of each domain described below. Gag is formed of four major domains (MA – Matrix, CA – Capsid, NC – Nucleocapsid, and p6) joined by two linker regions (SP1 and SP2). The specific functions encoded in Gag's domains coordinate key stages throughout the virus lifecycle, including Gag assembly, genome packaging and recruitment of ESCRT proteins to allow virions to bud from the host cell.
Source: Modified from (Freed, 2015).

Env is expressed as a single gp160 protein that is cleaved into its active gp120 and gp41 subunits by host proteases. The regulatory proteins are *trans*-activator of transcription (Tat) and regulator of expression of virion proteins (Rev), and the accessory proteins are viral protein U (Vpu), viral protein R (Vpr), virion infectivity factor (Vif) and negative factor (Nef).

In addition to these, key steps of the virus replication cycle are regulated by ribonucleic acid (RNA) structures in the genomic and subgenomic viral RNA (vRNA) species. A number of these are found at the 5' and 3' UTRs. The 5'UTR is composed of a repeat (R) region, a unique 5' region (U5) and a series of helix loop RNA structures termed stem-loops that perform critical functions in viral replication (Figure 3).

R is present at both the 5' and 3' ends of the genome and includes the *trans*-activation response element (TAR) and the poly(A) hairpin, although the former acts at the 5' end and the latter at the 3' end. TAR is recognised by the Tat protein to promote efficient transcription (Rosen et al., 1985). The 5' version is dominant although the mechanism underlying this imbalance is unknown (Klaver and

Berkhout, 1994). Downstream of R is U5, which harbours the primer binding site (Pbs), an 18 nucleotide sequence complementary to the 18 nucleotides at the 3' terminus of the host transfer RNA (tRNA) $tRNA^{Lys3}$ (Jiang et al., 1993).

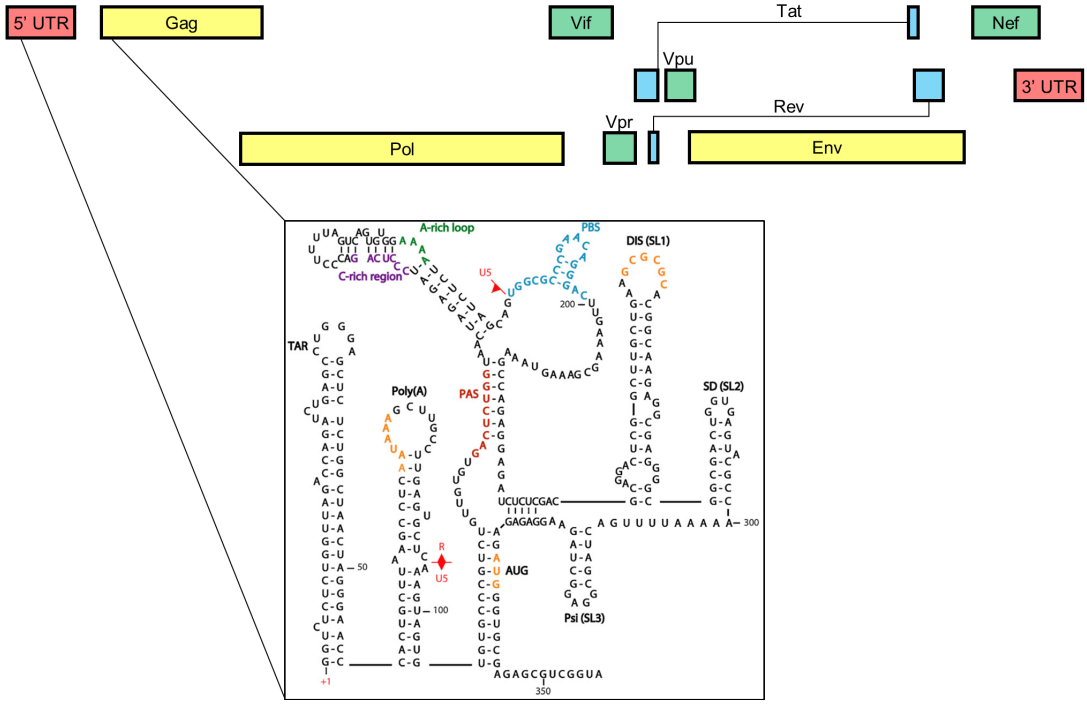
In between U5 and extending into the Gag open reading frame (ORF) are four stem-loops (Figure 3). Stem-loop 1 (SL1) is important for genome dimerisation and contains the 6-nucleotide DIS palindrome in its apical loop (Marquet et al., 1994; Muriaux et al., 1995; Paillart et al., 1994; Skripkin et al., 1994). It also contains Gag binding sites (Abd El-Wahab et al., 2014; Damgaard et al., 1998; Kenyon et al., 2015; Wilkinson et al., 2008) making it part of the core packaging signal (psi/Ψ) with SL3 being the other major component (Kim et al., 1994; Lever et al., 1989; McBride and Panganiban, 1996).

In between SL1 and SL3 is SL2, the major splice donor (Arrigo et al., 1990). As psi overlaps the splice donor site it is only found in full-length gRNA and not in fully- or partially-spliced vRNA. This packaging specificity is important as only viruses containing full-length gRNA encode all viral proteins and are replication competent. SL4 contains the AUG start codon for Gag. The four stem-loops are not rigid conformations but can adopt different structures, each of which have specific functions in the viral lifecycle (Berkhout and van Wamel, 2000; Kenyon et al., 2013; Lu et al., 2011).

Functional RNA structures also exist within genes. The Gag ORF contains a frameshifting signal (Jacks et al., 1988) composed of a slippery site and an RNA stem-loop. Slipping of ribosomes during translation enables -1 frameshifting to occur at a rate of 1/20, producing the GagPol fusion protein. Disturbance of this frequency is detrimental to virus infectivity (Shehu-Xhilaga and Crowe, 2001). In the coding region of Env is the Rev-response element (RRE), a 350 nucleotide *cis*-acting structure to which Rev binds to facilitate export of unspliced and partially-spliced vRNA from the nucleus (Rosen et al., 1988).

The poly(A) hairpin acts only at the 3' end of the genome due to the presence of upstream enhancer sequences in the unique 3' region (U3), where it stimulates polyadenylation of transcribed RNAs (Valsamakis et al., 1991).

Figure 3 – HIV-1 genome organisation and structures in the 5'UTR



*Overview of genome structure. Structural genes are highlighted in yellow, regulatory genes in blue and accessory genes in green. The UTRs are highlighted in red. Inset – Secondary structural depiction of the 5'UTR with key motifs highlighted, modified from (Sleiman et al., 2012). At the 5' end is the R region containing the TAR and Poly(A) hairpins. Downstream is the U5 region, containing the Pbs hairpin. This is followed by 4 stem-loops. SL1 contains the DIS, SL2 is the major splice donor, and SL3 is the major packaging signal. SL4 contains the AUG Gag start codon; in this structure, a stem-loop is not present and AUG is instead annealed to U5, as the 5'UTR exists in two conformations (**Figure 6**).*

1.4. Virion structure

In common with all retroviruses, HIV-1's two copies of its single-stranded ribonucleic acid (ssRNA) genome are packaged as a non-covalently linked dimer into the virion (Darlix et al., 1990; Paillart et al., 2004). Packaging a dimeric genome offers the capacity to switch template strands during reverse transcription. This has two clear benefits for viral replication and evolution. Firstly, given the fragile nature of ssRNA, the presence of two templates increases the probability of successful reverse transcription in the case of strand breaks (Coffin, 1979). Secondly, co-packaging of genetically distinct gRNAs in coinfected cells followed by template-switching in the target cell enables recombination to generate chimeric proviruses, a process which is a major driver of viral diversity (Burke, 1997; Hu and Temin, 1990). Recombination is extremely common: over 40 circulating recombinant viruses have been identified one of which (CRF01) was responsible for an epidemic in Thailand (Taylor et al., 2008).

In the mature virus particle the gRNA, coated with nucleocapsid (NC) proteins, is enclosed in a conical-shaped core composed of 1,200-1,500 capsid (CA) protein subunits (Briggs et al., 2003) derived from proteolytic processing of the Gag protein (Bell and Lever, 2013).

In addition to gRNA, the mature virion contains the viral enzymes necessary for initiating new rounds of infection. Reverse transcriptase (RT) has deoxyribonucleic acid (DNA) polymerase and RNase H activity, enabling it to catalyse production of a double-stranded DNA copy of the viral genome. Integrase (IN) facilitates insertion of this proviral DNA into the host cell genome to allow the virus to establish residency for the lifetime of the cell.

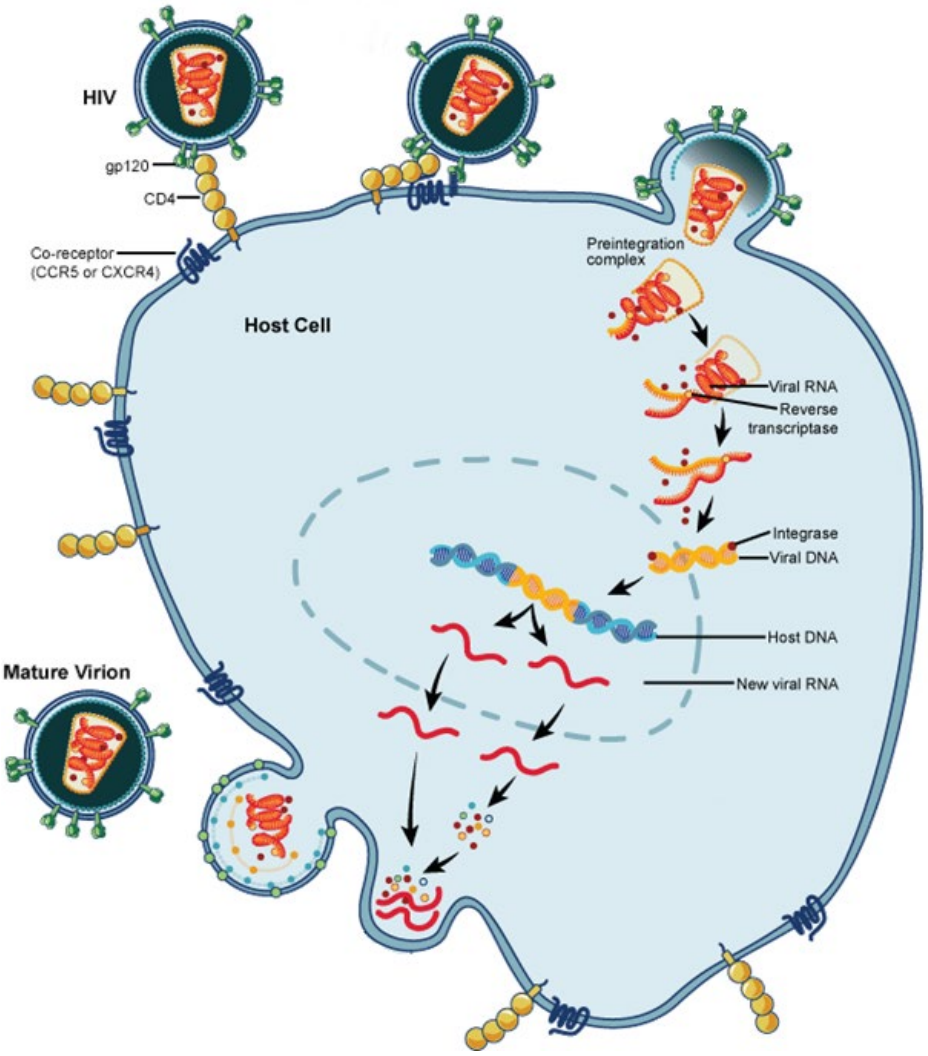
The mature conical core of the virus is enclosed in a lipid bilayer membrane derived from the host cell during budding. Embedded in this membrane are matrix (MA) proteins derived from processing of Gag, and Env glycoproteins. Env is required for a specific interaction with receptors on target cells and triggering of the subsequent fusion event. It is a class I fusion protein, being predominantly α-helical, and is expressed on the virion surface as a trimer of heterodimers. Each heterodimer consists of a surface-exposed gp120 subunit and a transmembrane gp41 subunit (Weiss et al., 1990). The former has receptor and coreceptor binding activity, and the latter contains the fusion peptide.

1.5. Virus entry

HIV-1 replicates in CD4 (cluster of differentiation 4)-positive helper T-cells (Dalgleish et al., 1984; Klatzmann et al., 1984), as well as monocytes, macrophages and dendritic cells, which express CD4, the receptor utilised by Env for binding and fusion (Maddon et al., 1986; McDougal et al., 1986). The virus also requires one of two coreceptors, C-C chemokine receptor 5 (CCR5) or C-X-C chemokine receptor 4 (CXCR4), binding to which is mediated by a conformational change triggered by binding of gp120 to CD4 (Figure 4). Coreceptor engagement triggers membrane fusion by exposing the gp41 subunit of Env that inserts into the host cell membrane. In this conformation, it is in a metastable “spring-loaded” state. The energy released by the subsequent folding back of the subunit brings the two membranes together to form the fusion pore (Melikyan, 2008).

Viruses that utilise the CCR5 coreceptor are termed R5-tropic, whilst those that utilise CXCR4 are X4-tropic. For reasons that are not entirely clear, transmitted founder viruses are almost exclusively R5-tropic (Keele et al., 2008). Over the course of infection the viral population in a proportion of infected individuals may shift towards an X4-tropic phenotype, which is associated with disease progression (Tersmette et al., 1989).

Figure 4 – HIV-1 replication cycle



Overview of HIV replication cycle from fusion to formation of mature virions. Virus particles fuse to the host cell membrane following interaction of gp120 with CD4 and a coreceptor (CCR5 or CXCR4). This releases the preintegration complex, containing viral RNA, reverse transcriptase, and integrase, into the cytoplasm. Reverse transcription leads to the formation of double-stranded viral DNA, which is transported across the nucleus and integrates into the host DNA, acting as a template for transcription of new viral RNA. Unspliced viral RNA (gRNA) serves two roles – as a template for translation of full-length Gag and Gag-Pol, and as the genome itself to be packaged. Gag-gRNA complexes embed in the plasma membrane and the newly formed virions hijack the host's ESCRT machinery to complete the budding process. Following budding, a series of coordinated cleavage events transform the virus from its immature form to an infectious mature one.

Source: modified from NIAID.

The predominant transmission of R5 viruses means that individuals homozygous for a CCR5 gene mutation termed CCR5Δ32, which prevents cell surface CCR5 expression, are resistant to HIV-1 infection. Heterozygous individuals have a reduced risk of infection and exhibit less severe disease

progression (Dean et al., 1996). This is exemplified by the case of the Berlin patient – a stem cell transplant from a CCR5Δ32 homozygous donor eliminated the virus from the patient in what remains the only known case of a cure (Hutter et al., 2009).

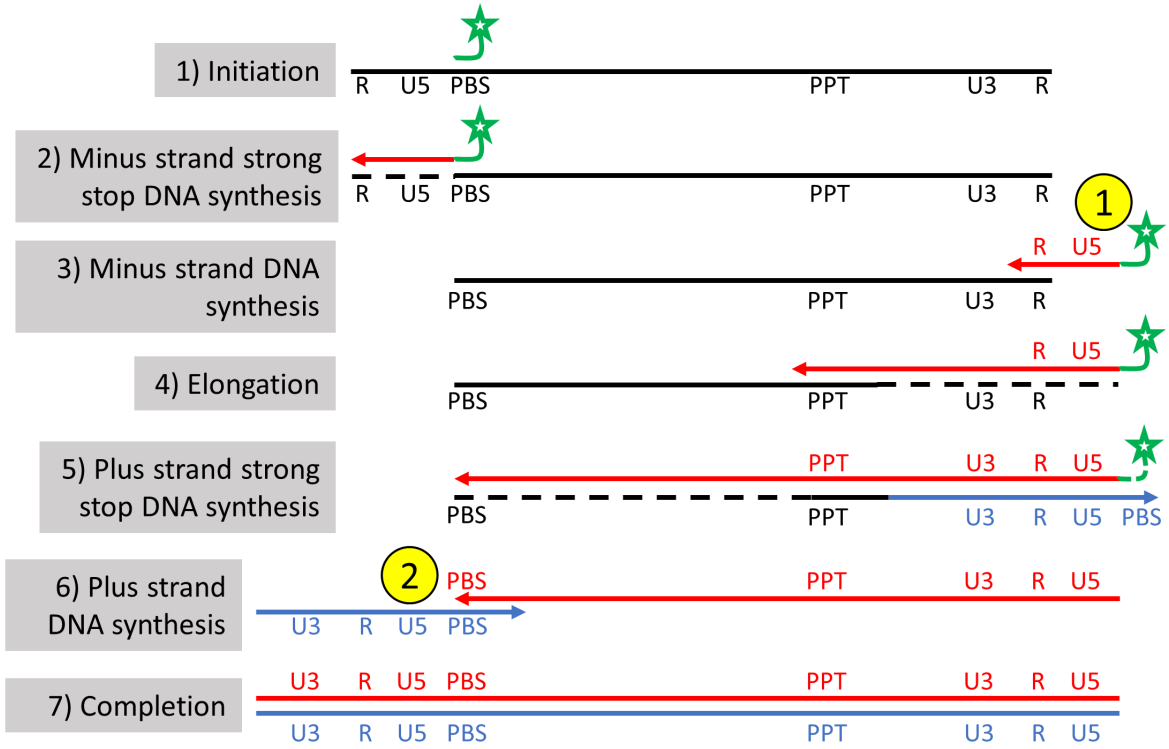
Aside from direct blood contact, HIV transmission between individuals occurs at the mucous membranes, which are heavily patrolled by macrophages and dendritic cells. Dendritic cells express low levels of CD4 so unlike CD4+ T cells they are not easily infected. However as antigen-presenting cells they can play an important role in dissemination of infection by trapping virions and transporting them to the lymph nodes where antigen is presented to T cells (Piguet and Steinman, 2007).

1.6. Reverse transcription, uncoating and nuclear import

Following entry of the viral capsid into the cytoplasm of the host cells, several key events occur. The single-stranded gRNA undergoes reverse transcription to produce a double-stranded DNA copy. The mature capsid dissociates in a process termed uncoating, which is required to enable the virus genome to traverse the nuclear pore complex (NPC), which is smaller in diameter (30 nm) than the capsid (50-60 nm) (Hilditch and Towers, 2014). This ability to cross the nuclear membrane permits lentiviruses to infect non-dividing cells, in contrast with other retroviruses which rely on the breakdown of the nuclear membrane during mitosis (Katz et al., 2005). The degree of overlap between the timing of these processes and the mechanisms linking them are incompletely understood.

The process of reverse transcription is illustrated in Figure 5. All retroviruses use host tRNAs as primers for reverse transcription (Mak and Kleiman, 1997). HIV-1 uses the $tRNA^{Lys3}$ primer, which primes the initial elongation of the minus strand from the Pbs towards the 5' end of the template gRNA. Meanwhile RNA in the nascent RNA:DNA duplex is degraded due to the RNase H activity of the RT enzyme (Step 1 and 2 in Figure 5). This 'strong stop DNA' is then translocated to the 3' end (first strand transfer) of either the same strand or the partner strand where it binds to the complementary R region. Elongation continues to produce the minus strand with simultaneous degradation of the template RNA (Steps 3 and 4 in Figure 5). Two polypurine tracts that are resistant to degradation serve as primers for plus strand synthesis. This is followed by second strand transfer and completion of plus strand synthesis, generating a double-stranded DNA molecule that is longer than the parental gRNA with U3, R and U5 long terminal repeat (LTR) sequences at each end (Steps 5–7 in Figure 5) (Hu and Hughes, 2012). The product of reverse transcription is a preintegration complex (PIC) containing blunt-ended double-stranded DNA and IN (Figure 4).

Figure 5 – HIV reverse transcription



Schematic illustrating the process of reverse transcription, described in the text above with reference to the steps in this diagram. Black lines represent viral RNA, and red and blue lines represent nascent proviral DNA strands. Dashed lines indicate degradation of RNA by RNase H activity. The host tRNA used to prime reverse transcription is coloured green. First and second strand transfers are highlighted by circled numbers.

Inhibition of reverse transcription delays uncoating, but most reverse transcription is completed after the onset of uncoating, suggesting that uncoating is triggered during reverse transcription (Hulme et al., 2011). More recent data suggests that the trigger for uncoating is the completion of first strand transfer (Cosnefroy et al., 2016; Mamede et al., 2017). Imaging of virions *in vitro* using atomic force microscopy has shown that that the build-up of pressure caused by reverse transcription causes virions to swell and disassemble (Rankovic et al., 2017).

Virus infectivity is sensitive to disturbances in the timing of uncoating, because the capsid is thought to conceal viral DNA from cytoplasmic DNA sensors, preventing stimulation of type I interferon (IFN) production (Manel et al., 2010; Rasaiyaah et al., 2013). The host protein cyclophilin A (CypA), which binds to CA (Luban et al., 1993) appears to be important for preventing premature uncoating; *in vitro* it stabilises the capsid (Shah et al., 2013). This protects the viral genome from cytoplasmic sensors, as inhibition of binding to CA reduces the levels of viral DNA (Braaten et al., 1996) and triggers an innate immune response (Rasaiyaah et al., 2013). Interestingly, Manel *et al* observed that when resistance to HIV-1 is subverted in dendritic cells (dendritic cells are normally resistant to HIV-1 infection) an

interaction between CypA and newly synthesised Gag triggers innate immune sensing through a cryptic mechanism. This highlights an evolutionary balancing act whereby CA's interaction with CypA prevents premature uncoating and boosts infectivity, whilst the interaction between newly synthesised CA and CypA triggers innate immune sensors (Manel et al., 2010). It is unlikely that the entire capsid disintegrates before the genome reaches the NPC, as this would presumably trigger cytoplasmic DNA sensing. This is supported by the finding that CA can directly interact with the NPC protein Nup358 (Di Nunzio et al., 2012; Schaller et al., 2011). Furthermore, CA can be visualised in the nucleus (Chin et al., 2015) where it appears to play a role in integration (Chen et al., 2016).

Some have proposed that a fully intact capsid docks at the nuclear pore (Arhel et al., 2007; Matreyek and Engelman, 2011; Schaller et al., 2011). However, a recent study using live-cell fluorescent imaging revealed that uncoating takes only 30 minutes to occur following fusion and CA's roles in later steps are facilitated by a partial capsid reaching the nucleus (Mamede et al., 2017).

Another protein which binds to CA is cleavage and polyadenylation specificity factor-6 (CPSF6). Disruption of this interaction delays reverse transcription and uncoating (De Iaco et al., 2013; Fricke et al., 2013) further highlighting the link between these two processes.

The details of nuclear localisation of the PIC have not been fully established, but short interfering RNA (siRNA) screens have identified roles for a transport protein, TNPO3, and NPC proteins Nup153 and Nup358 in infection (Brass et al., 2008; König et al., 2008). Disruption of CypA or CPSF6 binding to CA relieves the requirement for Nup358 and TNPO3 perhaps by directing the PIC through alternative nuclear import pathways (Lee et al., 2010; Schaller et al., 2011).

It appears that CypA and CPSF6 binding to CA are important for delaying the onset of reverse transcription and uncoating, in addition to directing the PIC along a nuclear import pathway using TNPO3, Nup358 and Nup153 (Hilditch and Towers, 2014).

1.7. Integration and latency

Prior to the use of next generation sequencing, analysis of sites in the genome into which HIV integrates was limited, but the topology of the target DNA was found to be important; integration is favoured where DNA is bent or distorted (Pruss and Wolffe, 1994). New sequencing technologies allowed more thorough investigation, and it was found that integration is strongly favoured within genes and in particular within those which are highly actively transcribed (Schröder et al., 2002). IN seems to be the main determinant of integration site selection, since swapping the HIV IN enzyme for the murine leukemia virus (MLV) equivalent changes integration site preference to sites favoured by MLV (Lewinski et al., 2006). A host protein, LEDGF, was also shown to be important in target site selection; it binds to IN (Cherepanov et al., 2003) and knockdown of the protein reduces preferential targeting for active transcriptional units (Ciuffi et al., 2005). LEDGF knockdown did not block

integration in Ciuffi *et al*, however a later study using more intense knockdown and a dominant-negative mutant demonstrated that LEDGF is in fact essential for integration (Llano et al., 2006). This discrepancy is due to the presence of a small but highly potent pool of chromatin-associated LEDGF remaining in the cell when less intense knockdown methods are used. Integration is linked to the preceding steps of virus replication as mutations in CA preventing CypA or CPSF6 binding alter integration site preference (Schaller et al., 2011).

Once at the integration target site IN catalyses the reaction steps required for the viral DNA to be inserted into the host DNA. The two 3'-most nucleotides of each strand of the blunt-ended DNA are first removed, and the new 3' terminal bases of the viral DNA attack the phosphodiester bonds in the host DNA integration site, before becoming covalently attached. Host repair enzymes then fill in the gaps to leave the provirus integrated (Figure 4) with a duplicated sequence at each end (Craigie and Bushman, 2012).

Once integrated the proviral DNA can be used as a template for transcription by the host machinery, creating new virus particles (Figure 4). Alternatively, it can be silenced by epigenetic changes or by transcriptional interference from neighbouring transcripts, causing the virus to establish latency (Siliciano and Greene, 2011).

1.8. Gene expression

Production of new vRNA species from the integrated provirus is performed by the host's transcription and splicing machinery. The viral LTR is an efficient promoter, constituting three Sp1 binding sites, a TATA box and an initiator sequence (Rittner et al., 1995), facilitating recruitment of transcription factor II D, a component of the RNA polymerase II preinitiation complex. Transcription efficiency is further increased by an enhancer element, composed of two NF-kB binding sites (Nabel and Baltimore, 1987), which is essential for reactivation of latent proviruses in CD4+ T-cells (Alcamí et al., 1995).

However, recruitment of host factors alone is not sufficient for efficient transcription and transport of vRNA to the cytoplasm; two viral factors – Tat and Rev – are critical for these processes to occur (reviewed in Karn and Martin Stoltzfus, 2012).

Tat, which is expressed from the central portion of the viral genome (Figure 3), is essential for viral transcription (Sodroski et al., 1985). Similar trans-activating factors are also found in bovine leukaemia virus and human T-cell lymphotropic virus (HTLV) types I and II (Sodroski et al., 1984). In the absence of Tat transcription initiation is unaffected, but transcription complexes stall shortly after initiation, demonstrating that Tat functions to promote efficient elongation of viral transcripts (Kao et al., 1987). It binds to a U-rich bulge near the tip of the TAR stem--loop (Dingwall et al., 1989), a in the 5'UTR (Figure 3), in complex with the CDK9 kinase component of the elongation factor pTEFb (Herrmann and Rice, 1995) and Cyclin T1 (Wei et al., 1998). Binding of Tat to CDK9 results induces a conformational

change in the enzyme, activating pTEFb resulting in phosphorylation of elongation factors and the C-terminal domain of RNA polymerase II (Isel and Karn, 1999).

The viral genome contains multiple splice donor and splice acceptor sites giving rise to diverse vRNA species through alternative splicing (Purcell and Martin, 1993). Unspliced and partially-spliced RNAs are unable to exit the nucleus unaided because the cell has mechanisms in place to prevent intron-containing RNA from being exported and thus translated (Cullen, 2000).

To overcome this problem, the viral Rev protein binds to the RRE in unspliced and partially-spliced transcripts, facilitating their export from the nucleus to the cytoplasm (Rosen et al., 1988; Sodroski et al., 1986). As the RRE is in the *Env* coding region, it is only present in incompletely spliced RNAs.

Rev multimers bound to the RRE interact with the transport factor Crm1 through a nuclear export signal (Neville et al., 1997), enabling the complex to be transported through the NPC and into the cytoplasm. Similar regulatory factors are also produced by HTLV I & II (Romanelli et al., 2013). By contrast, the genome of the simple retrovirus Mason-Pfizer monkey virus (M-PMV) contains an RNA structure termed a constitutive transport element that interacts directly with RNA export machinery (Cullen, 2003).

Rev, Tat and Nef are described as 'early' genes because they are translated from fully-spliced transcripts and can therefore be expressed in a Rev-independent manner. Once sufficient Rev has accumulated unspliced and partially-spliced transcripts can exit the nucleus to enable translation of incompletely spliced 'late' genes.

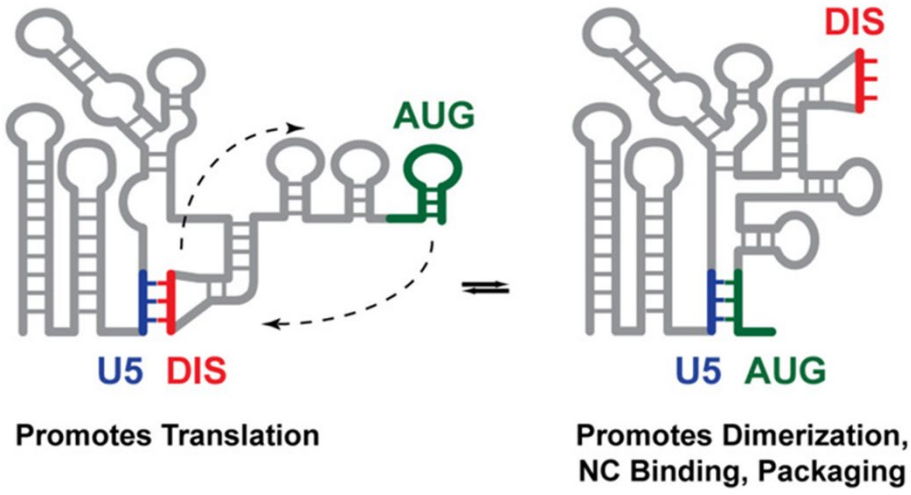
1.9. Structural switch in the 5'UTR regulating translation and packaging

The unspliced RNA serves as both the gRNA to be packaged and as a template for Gag translation. These are two mutually exclusive processes. 5'UTR transcripts migrate with two distinct mobilities in gels suggesting the presence of two conformers (Berkhout and van Wamel, 2000). Mutational analysis showed that formation of the faster-migrating conformer required the poly(A) and DIS regions (Huthoff and Berkhout, 2001). Initial studies using RNA structure prediction tools proposed that in the fast migrating structure SL1 completely melts and anneals with the U5 region to form an elongated stem (Abbink and Berkhout, 2003; Huthoff and Berkhout, 2001). However, more recent data from nuclear magnetic resonance (NMR) spectroscopy and selective 2'-hydroxyl acylation analysed by primer extension (SHAPE) probing have refined this model. The DIS palindrome at the tip of SL1 anneals to residues in U5 forming a pseudoknot (Figure 6 left) (Kenyon et al., 2013; Lu et al., 2011).

The two conformers perform different functions - packaging and translation - and the structural switch regulates the balance between these. In one conformation SL1 is occluded by base pairing between the DIS and U5, preventing dimerisation but promoting translation. In the other, SL1 is exposed to

enable dimerisation but translation is inhibited due to the U5-AUG interaction (Figure 6). Packaging is promoted in the latter conformer since this arrangement of SL1, SL3 and sequences in the 5' sequence of *gag* form the structure recognised by the NC domain of Gag (Section 1.12) (Kenyon et al., 2015; Kutluay et al., 2014; Lu et al., 2011).

Figure 6 – Structural switch in the 5'UTR



The 5'UTR can adopt two conformations, promoting either translation (left), or dimerisation and packaging (right). SL1 contains the DIS palindrome at its tip (highlighted in red). Downstream of this are SL2, SL3 and SL4, which contains the Gag start codon (highlighted in green).
Source: (Lu et al., 2011).

Examples of structural switches controlling the balance between translation and packaging are also found in feline immunodeficiency virus (FIV) (Kenyon et al., 2011) and the more distantly related Moloney murine leukemia virus (MoMuLV) (Miyazaki et al., 2010; Mougel et al., 1993).

In vitro, formation of the pseudoknot conformer which promotes packaging can be triggered by binding of the NC protein (Huthoff and Berkhout, 2001). *In vivo*, isolated NC protein is not available at this stage and gRNA is believed to be initially captured by a small number of full-length Gag molecules (Hendrix et al., 2015; Mailler et al., 2016). Nevertheless the unprocessed NC domain within Gag likely performs the same function, and this is consistent with the observation that at low concentrations Gag stimulates translation, whilst at a higher concentrations it is inhibitory (Anderson and Lever, 2006). Together, these data suggest that once sufficient Gag has accumulated for virion assembly the 5'UTR changes from a molecule promoting translation to one promoting dimerisation and packaging.

Work from the Summers lab has led to an alternative hypothesis. They propose that the structural and therefore functional fate of each gRNA is determined at the moment of transcription, rather than a situation whereby each gRNA possesses the ability to switch between the two functions. Heterogeneity in transcription start site usage was shown to result in different numbers of G residues

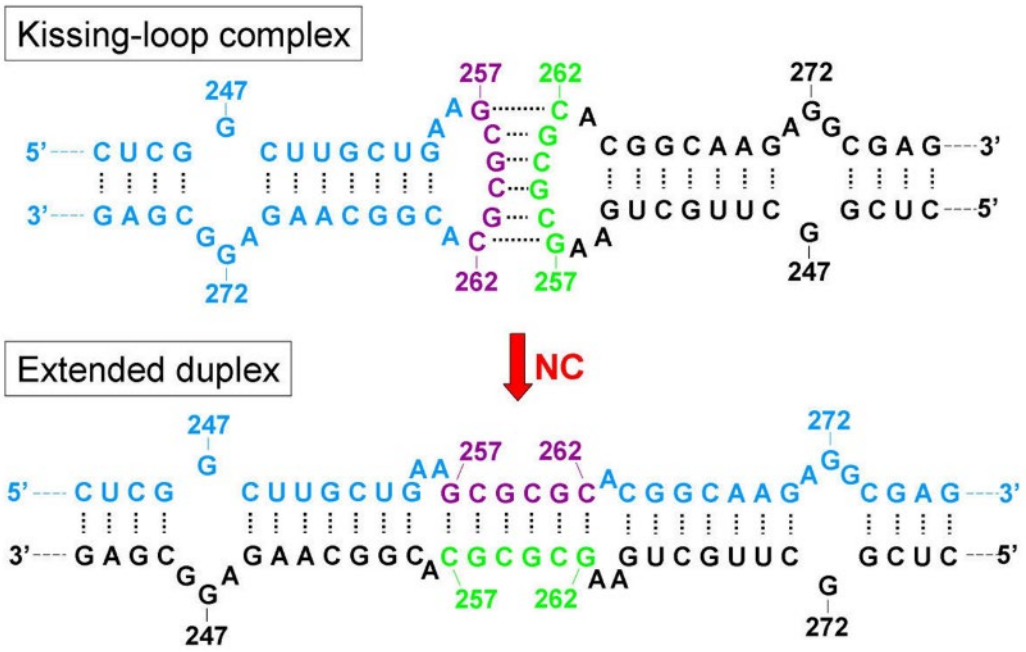
at the 5’ end of gRNA. gRNAs with a single G residue at the 5’ end formed dimers and were preferentially packaged into virions, whilst those with two or three G residues formed monomers and were enriched on polysomes. It was proposed that the additional G residues destabilise the poly(A) hairpin, promoting the U5:DIS interaction (Figure 6 left) (Kharytonchyk et al., 2016).

1.10. gRNA dimerisation

Genome dimerisation is a unique feature of retroviruses (Paillart et al., 2004). When viewed under an electron microscope, RNA extracted from HIV-1 virions is seen as a dimer connected at the 5’ end of the genome (Höglund et al., 1997). Early studies on dimerisation of *in vitro* transcribed RNA suggested that this dimer linkage structure (DLS) was formed by contacts between polypurine tracts downstream of the 5’ major splice donor (Darlix et al., 1990; Marquet et al., 1991). As these sequences are found only in full-length gRNA this seemed a plausible mechanism for ensuring that only full-length gRNA can dimerise. The DLS was later found to be an artefact of using short transcripts *in vitro*, and is not important for dimerisation *in vivo* (Haddrick et al., 1996), however the DIS sequence upstream of the splice donor was found to be critical.

The DIS is a 6-nucleotide palindrome in the apical loop of SL1. Dimerisation of gRNA is facilitated by intermolecular base pairing between DIS sequences (Haddrick et al., 1996; Marquet et al., 1994; Muriaux et al., 1995; Paillart et al., 1994; Skripkin et al., 1994). During virion maturation, the weak kissing-loop dimer is converted to a stronger extended dimer when the stems of SL1 melt and anneal to complementary sequences on the partner gRNA molecule (Figure 7) (Skripkin et al., 1994).

Figure 7 – Maturation of the gRNA dimer



The initial kissing-loop interaction formed by base pairing of DIS nucleotides is converted to a strong extended dimer interaction during virion maturation. NC catalyses the unwinding of SL1 and the establishment of extensive inter-strand base pairing.

Source: (Russell et al., 2004).

The importance of the DIS in initiating dimerisation explains the 100% conservation of palindromic sequences across all tested HIV-1 isolates (Berkhout, 1996). The palindromic nature of the DIS loop rather than its sequence appears to be critical for dimerisation. However some non-native palindrome sequences do severely impact dimerisation and virus replication, so sequence constraints do exist (Laughrea et al., 1999).

Sequences involved in dimerisation and packaging overlap or are in close proximity to each other suggesting a link between the two processes (Abd El-Wahab et al., 2014; Clever and Parslow, 1997; Harrison et al., 1998; Kim et al., 1994; Laughrea et al., 1999, 1997; Lever et al., 1989; Liang et al., 1998; McBride and Panganiban, 1997, 1996; Paillart et al., 1996; Sakuragi et al., 2003; Shen et al., 2001, 2000).

There has been extensive debate about whether dimerisation precedes packaging or *vice versa* (Russell et al., 2004), but convincing evidence now exists demonstrating that formation of a dimer enables recognition by Gag for packaging. Duplication of the DIS in the genome results in the packaging of monomeric RNA. This indicates that an intra-molecular dimer can substitute for an inter-molecular one for recognition by Gag, showing that it is the unique structure of the dimer that is recognised for packaging (Nikolaitchik et al., 2013; Sakuragi et al., 2002, 2001). Furthermore, recombination between

subtype B (GCGCGC palindrome) and C (GUGCAC palindrome) viruses is less frequent than intra-subtype recombination; a barrier which can be overcome by mutating their palindrome sequences to match (Chin et al., 2005; Moore et al., 2007). For recombination to occur heterologous RNAs must first be co-packaged, thus discordant palindromes, which are unable to form dimers, are less likely to be packaged than concordant ones.

In recent years, it has become possible to study dimerisation taking place in cells using advanced microscopy techniques such as super resolution microscopy, and total internal reflection fluorescence microscopy (TIRFM), which allows visualisation of molecules at or near the plasma membrane. gRNA can be labelled by introducing sequences encoding RNA stem-loops recognised by MS2 from bacteriophage MS2 or BglG from *Escherichia coli* (*E. coli*) into the genome and co-expressing with a fluorescently labelled version of the stem-loop's cognate binding protein.

Chen *et al* developed the technique of co-expressing gRNAs labelled with two distinct fluorescent proteins (Chen et al., 2009). Measurement of the proportion of dual-labelled foci enables a direct determination of dimerisation efficiency in cells and virions. If only monomeric gRNA is present, then all foci would be either one colour or the other. However, if the gRNA population consists entirely of dimers then 50% of foci would be dual-labelled. The proportion of dual-labelled RNA in virions is close to 50% supporting the model whereby dimers are specifically selected for packaging (Chen et al., 2009).

When TIRFM is used, gRNA is seen to move dynamically in and out of the membrane with a single step change in fluorescence, suggesting the arrival of pre-formed dimers at the membrane (Jouvenet et al., 2009). This is supported by observation that HIV-1 gRNA colocalisation is 6-fold higher than non-viral RNA colocalisation in the cytoplasm (Ferrer et al., 2016). However, dimers were more enriched at the plasma membrane suggesting selective targeting of dimers to the plasma membrane and/or *de novo* initiation of dimerisation at the plasma membrane. This is consistent with the observation of signals from gRNA molecules merging on the membrane (Chen et al., 2015).

Dimerisation is stimulated by Gag, as Gag expression increases the formation of dimers on the plasma membrane (Chen et al., 2015), and in the absence of Gag, dimerisation taking place in the plasma membrane and cytoplasm is 5-fold less efficient (Ferrer et al., 2016). Intriguingly, an NC mutant lacking gRNA binding was able to moderately stimulate dimerisation in the cytoplasm but not at the plasma membrane, suggesting that NC domain binding to gRNA is important for dimer accumulation at the plasma membrane (Ferrer et al., 2016).

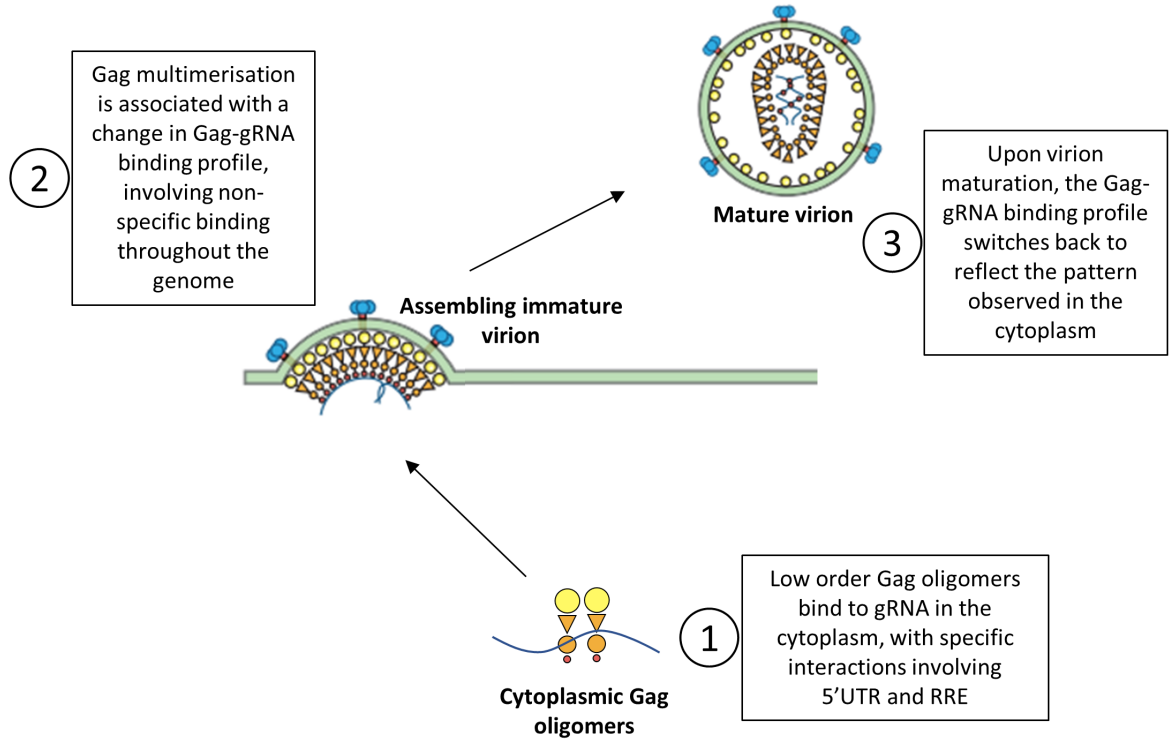
These results suggest that Gag acts as a chaperone to stimulate dimerisation or to stabilise newly formed dimers, however this does not contradict the finding that dimerisation precedes packaging. In agreement with biochemical and genetic data, heterodimerisation between individual gRNA molecules with discordant DIS sequences was severely inhibited preventing formation of heterodimer-containing

virions (Ferrer et al., 2016), and in the majority of cases particle assembly only occurred once a dimer had formed (Chen et al., 2015; Jouvenet et al., 2009).There is therefore a distinction to be made between Gag's role in stimulating dimerisation and in packaging, which involves formation of whole virus particles around the genome.

1.11. Spatiotemporal dynamics of Gag-gRNA interactions

The notion of different modes of Gag binding to RNA, one involving cytoplasm-associated oligomers and the other membrane-associated multimers, is consistent with the finding by Kutluay *et al* that the sequences Gag associates with in the viral genome change during assembly and the formation of virions. Using CLIP-sequencing (CLIP-seq) they were able to probe the interactions between Gag and gRNA sequences in cells and virions. Cytoplasmic Gag associated with sequences around the packaging signal in the 5'UTR, but also with the RRE. However, in immature virions this pattern of binding was diminished, and instead Gag associated with numerous sequences A-rich sequences throughout the genome. This switch in binding appeared to be driven by Gag multimerisation, as a mutant in the carboxy-terminal domain of CA (CA_{CTD}) exhibited a binding profile reminiscent of cytoplasmic Gag. Upon virion maturation, the binding profile reverted to the one observed for cytoplasmic Gag, demonstrating that throughout the viral assembly process the interactions between Gag and gRNA are dynamic (Kutluay et al., 2014) (Figure 8).

Figure 8 – Changes in Gag-gRNA binding profile throughout virion assembly



The interaction between Gag and gRNA is not static, changing multiple times throughout the virion assembly process, as illustrated in Steps 1–3 in the diagram. Early microscopy work indicated that gRNA doesn't interact with Gag until it reaches the plasma membrane, but this was due to the inability to detect low order Gag oligomers. Instead, these oligomers are thought to interact with gRNA in the cytoplasm, with the 5'UTR and RRE as the main points of contact (Step 1). Upon multimerisation at the plasma membrane, these specific interactions are replaced by non-specific binding throughout the genome (Step 2). Gag-gRNA interactions in mature virions resemble that of those in the cytoplasm, with binding specificity returned to the 5'UTR and RRE (Step 3).

Source: Modified from (Sundquist and Kräusslich, 2012)

Advances in imaging techniques have provided further evidence about the dynamics of the interaction between Gag and gRNA. Gag can be labelled by introducing sequences encoding fluorescent domains such as GFP or mCherry into its ORF and co-expressing with gRNA labelled using the methods described above.

Tracing individual gRNA molecules revealed that they move in the cytoplasm in a random-walk motion suggestive of diffusion, both in the absence and presence of Gag, and it was concluded that Gag does not interact with gRNA until it reaches the plasma membrane (Chen et al., 2014). However, gRNA is a much larger molecule, with a molecular weight 35-fold higher than Gag, so the effect of potential Gag binding on the movement of gRNA is likely to be minimal (Chen et al., 2014).

Using TIRFM it was observed that RNA signals appear at the plasma membrane before Gag signals, suggesting that RNA does not interact with Gag until it reaches the plasma membrane (Jouvenet et al., 2009). However, the sensitivity of detection in those experiments was 12 Gag molecules, so these results don't exclude the possibility that gRNA arrives at the membrane in complex with a smaller number of Gag molecules.

Hendrix *et al* overcame the Gag detection sensitivity issue by using fluorescence fluctuation spectroscopy, which enables inferences to be made regarding particle stoichiometry. Rather than looking at the effect of Gag on gRNA movement they investigated the effect of gRNA binding on Gag movement. Gag being a much smaller molecule makes it more likely that differences in motion will be observed if an interaction occurs. They found that Gag moves much more slowly in the cytoplasm than it should do based on its size, and that its mobility could be increased by mutating its RNA-binding NC domain. These data suggest that Gag and gRNA do indeed interact in the cytoplasm. NC mutation inhibited oligomerisation so they proposed that gRNA first interacts with Gag in the cytoplasm and nucleates assembly of a Gag oligomer. This is followed by anchoring in the plasma membrane where Gag multimerisation and particle assembly occurs (Hendrix et al., 2015). These findings are supported by crosslinking immunoprecipitation (CLIP) data showing that some cytoplasmic Gag dimers are associated with gRNA, with higher order Gag complexes only found at the plasma membrane (Kutluay and Bieniasz, 2010).

1.12. Genome packaging

HIV-1 genome packaging is remarkably efficient – over 90% of virions contain a gRNA dimer (Chen et al., 2009). The *cis*-acting determinants of packaging are found in the 5'UTR, and of particular importance are SL3 and SL1, as initially determined by reverse genetics experiments (Kim et al., 1994; Lever et al., 1989; McBride and Panganiban, 1996). Disruption of other regions also affects packaging, including SL4 (McBride and Panganiban, 1997) the TAR hairpin (Clever et al., 1999), poly(A) hairpin (Das et al., 1997) and U5-Pbs region (Russell et al., 2002). This may be due to secondary effects on the structure of the 5'UTR.

The stem-loops do not function in isolation but form part of a much larger structure (Figure 3), so manipulation of individual stem-loops can have unpredictable effects on the overall structure of the region. Structural studies have helped to put these observations into context.

A variety of structures have been predicted for the 5'UTR, reflecting the conformational flexibility of this region. However there is good agreement between structures derived using NMR (Lu et al., 2011), small-angle X-ray scattering (Jones et al., 2014) and single-molecule fluorescence resonance energy transfer (Stephenson et al., 2013), showing U5 bound to the start codon of Gag (Figure 3). Another NMR-derived structure shows the region comprising U5 and SL1 to SL4 folding into a 3-way junction

structure with SL2 being integral to the junction structure rather than forming a hairpin (Keane et al., 2015).

The *trans*-acting determinant of packaging interacting with the aforementioned *cis*-acting RNA structures is the NC domain of Gag (Figure 2). The key structural components of NC that enable specific packaging of gRNA are its two zinc finger motifs, each with the completely conserved sequence $CX_2CX_4HX_4C$ (where C = cysteine, H = histidine and X = any amino acid). In addition, NC contains basic residues that bind RNA non-specifically through electrostatic interactions. The spacer peptide 1 (SP1) domain immediately upstream of NC has also been shown to participate in packaging (Kaye and Lever, 1998).

In addition to reverse genetics and RNA structural probing, RNA protection and crosslinking assays have helped to map, at single-nucleotide resolution, the sequences and structures in the 5'UTR with which Gag interacts (Abd El-Wahab et al., 2014; Damgaard et al., 1998; Kenyon et al., 2015). These concur with data obtained using the above-mentioned techniques in that Gag binding sites map mostly to SL3 and SL1.

Despite being separated in sequence space, regions identified as being important for packaging and Gag binding are adjacent in the three-dimensional structure (Stephenson et al., 2013) and may be bound by the same Gag molecule. Indeed, mapping of at least 10 potential Gag binding sequences, determined by crosslinking and SHAPE, onto the three-dimensional structure of psi revealed that they make up four major Gag binding interfaces (Kenyon et al., 2015).

CLIP-seq data show that the region in the 5'UTR which Gag binds to in the cytoplasm comprises the sequences involved in the long-range interaction between U5 and AUG in addition to SL1 and SL3 (Kutluay et al., 2014). This pattern of Gag binding is more consistent with models showing SL2 adopting a hairpin structure (Figure 3) (Kenyon et al., 2013; Lu et al., 2011) instead of being involved in long-range base pairing (Keane et al., 2015). The findings of Kutluay et al also highlight another key determinant of successful genome packaging – the A-rich nature of the HIV-1 genome. In their CLIP-seq experiments they showed that Gag only modestly favours binding to gRNA over host RNAs in the cytosol, however upon multimerisation it becomes much more selective through a preference for A-rich regions in the gRNA over the GU-rich host RNAs (Kutluay et al., 2014).

This process does not however exclude packaging of host RNAs and spliced vRNAs, which make up around half of the RNA in the virion by mass (Eckwahl et al., 2016). Their inclusion is not random as demonstrated by the failure to detect other host RNAs such as GAPDH (Houzet et al., 2007). As described earlier, HIV-1 uses the host $tRNA^{Lys3}$ as a primer for reverse transcription through interaction with the Pbs, and other tRNAs interact with MA to prevent it from binding to internal membranes (Kutluay et al., 2014). Additionally, non-coding host RNAs such as U6 and 7SL – which are product of

RNA Pol III transcription – feature prominently in virus particles. Best-studied is the 7SL RNA, which is enriched 250-times more than actin RNA (Onafuwa-Nuga et al., 2006). One potential role for these non-viral RNAs is to assist in multimerisation of Gag through tethering of basic regions of NC (discussed further below). Another possibility is that they assist in host defence against HIV infection. There has been debate about the identity of the RNA that assists in packaging of the restriction factors APOBEC3G (Eckwahl et al., 2016) with one study finding 7SL RNA to be important for APOBEC3G incorporation (Wang et al., 2007). Definitive evidence is lacking at present, but it very unlikely that these host RNAs are enriched by chance.

1.13. Gag multimerisation and particle assembly

Once gRNA has reached the plasma membrane accompanied by oligomers of Gag, particle assembly begins (Figure 4) (Jouvenet et al., 2008; Kutluay and Bieniasz, 2010). Using microscopy, Gag-gRNA complexes are observed to stabilise on the membrane followed by increasing intensity of the Gag signal, indicating Gag multimerisation taking place (Chen et al., 2015; Jouvenet et al., 2009). As described above there is also a dramatic change in the binding of Gag to gRNA. Binding preference is switched from specific sites in the genome – psi, the RRE and 3'UTR – to less specific sites along the entire length of the genome (Kutluay et al., 2014).

Most other retroviruses also assemble their capsids at the plasma membrane, however some preassemble immature capsids within the cytoplasm. An historical system classifying retroviruses by virion morphology describes different subtypes: B-type (extracellular, eccentric, spherical core), C-type (central, spherical core) and D-type (cylindrical core) (Fenner, 2016). C-type retroviruses assemble their capsids at the plasma membrane during budding (similar to HIV) and include HTLV, and Avian leukosis and sarcoma viruses (Weiss, 1996). B- and D- type retroviruses, on the other hand, assemble their capsids in the cytoplasm before budding – it is even possible to alter the morphogenesis of the D-type retrovirus M-PMV, by introducing a point mutation in matrix, to resemble that of a C-type retrovirus (Rhee and Hunter, 1990). A Gag domain (p12) in M-PMV common to B- and D- type retroviruses (but absent in C-types) is important for intracytoplasmic assembly at lower (physiological) levels of Gag expression but, is redundant for cytoplasmic assembly at higher levels of expression (Sommerfelt et al., 1992).

Stabilisation and subsequent multimerisation of HIV Gag on the plasma membrane is dependent upon an intact MA domain (Figure 2) (Jouvenet et al., 2009; Kutluay and Bieniasz, 2010). At the N-terminus of MA is a myristic acid moiety, which enables it to embed into the plasma membrane (Göttlinger et al., 1989). This moiety is initially sequestered within a hydrophobic pocket in MA and is only exposed upon binding of the hydrophobic pocket to $PI(4,5)P_2$, a phospholipid found exclusively in the plasma membrane (Chukkapalli and Ono, 2011; Saad et al., 2006).

MA is also capable of binding to nucleic acid in addition to NC. Incubation of MA with PI(4,5)P_2-containing liposomes abolishes its RNA binding activity, demonstrating that RNA and plasma membrane lipids compete for binding to MA (Alfadhli et al., 2009). CLIP-seq data revealed that MA binds almost exclusively to a particular subset of tRNAs, and this interaction inhibits internal membrane binding because RNase treatment of cell lysates causes Gag to redistribute to membranes (Kutluay et al., 2014).

The mechanisms described above delay membrane binding and virion release, suggesting that there is a selective pressure for virus release to not occur too rapidly. Deletion of the globular head of MA enables faster and more abundant virus release, however this fails to give sufficient time for Vif to be expressed. Vif is a late gene which degrades the restriction factor APOBEC3G, which is packaged into virions where it mutates deoxycytidine residues to deoxyuridine in the nascent DNA strand during reverse transcription. By delaying release the virus has time to express Vif to maximise the likelihood of successful infection of new host cells (Holmes et al., 2015).

The NC domain of Gag is also important for indirect Gag-Gag interactions mediated by non-specific RNA binding. Unlike for the initial oligomerisation event in the cytoplasm, gRNA is not required for Gag multimerisation, as the rate of Gag multimerisation is independent of the presence of gRNA (Jouvenet et al., 2009). The assembly of CA-NC complexes *in vitro* is stimulated by the presence of non-viral RNA (Campbell and Vogt, 1995), and retroviral particles produced *in vivo* degrade when treated with RNase (Muriaux et al., 2001). Furthermore, mutation of basic residues in NC reduces the ability of Gag to multimerise *in vitro* and *in vivo* whereas zinc finger mutations have little effect (Cimarelli et al., 2000). NC's role in Gag multimerisation can be achieved by substituting it with a self-dimerising leucine zipper motif (Accola et al., 2000).

Virus-like particles (VLPs) containing Gag alone can be assembled *in vitro* (Campbell and Rein, 1999) and *in vivo* (Gheysen et al., 1989) in the absence of gRNA, and can also be released from cells. This appears to contradict the observation that gRNA nucleates particle assembly. However, gRNA may be important for nucleation under physiological conditions where intracellular Gag is less concentrated.

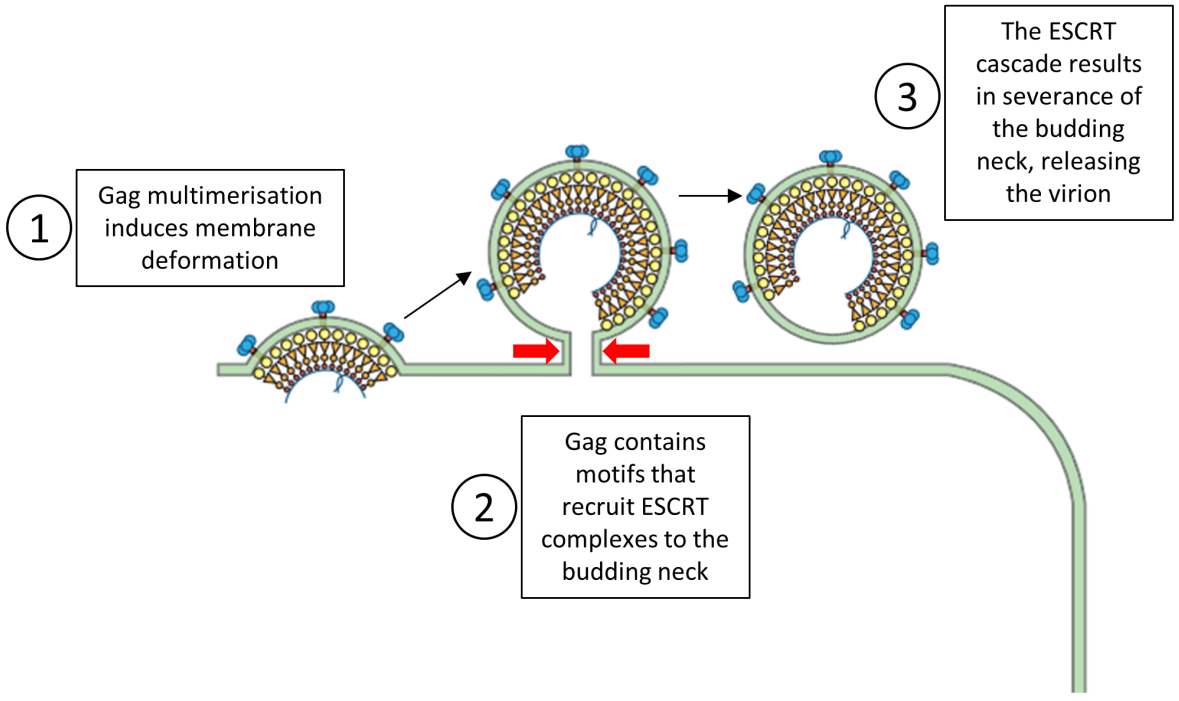
gRNA plays an important role in virus assembly in Rous sarcoma virus (RSV). RSV Gag takes a different route towards the plasma membrane, briefly travelling to the nucleus assisted by nuclear localisation signals (NLS) in NC and MA. It is in the nucleus that Gag packages gRNA before returning to complete assembly at the plasma membrane. Association of the packaging signal in gRNA facilitates nuclear export by helping to expose a nuclear export signal in the Gag p10 domain (Parent, 2011). In HIV-1, mutation of the gRNA-binding motif in NC causes Gag nuclear localisation, although the implications of this are unclear (Grigorov et al., 2007).

CA_{CTD} and the SP1 domain are particularly important for the establishment of direct contacts between adjacent Gag molecules (von Schwedler et al., 2003). Within CA_{CTD} is the major homology region, a highly conserved 20 amino acid sequence common to all retroviruses (Wills and Craven, 1991). In the immature virion Gag molecules are arranged in a hexameric lattice, with CA-CA interactions supported by 6-helix bundles formed by SP1 domains (Wright et al., 2007).

1.14. Virus budding

Following Gag multimerisation, virus particles must bud from the cell (Figure 4). Formation of a lattice induces deformation of the membrane to curve away from the cytoplasmic face, but the virus does not encode the apparatus to sever the budding neck. Instead, like a number of enveloped viruses including the Ebola virus (Garrus et al., 2001; Martin-Serrano et al., 2001), it hijacks the host's endosomal sorting complexes required for transport (ESCRT) machinery.

Figure 9 – Topology of HIV budding



The Gag protein alone is capable of assembling into virion structures (Step 1), but the virus lacks the machinery to catalyse the membrane severing event required to release the budding particle from the cell surface. Virus budding is topologically equivalent to the membrane remodelling events that take place during cytokinesis, and HIV and several other enveloped viruses take advantage of this similarity by hijacking the host cell's own ESCRT machinery (Step 2). The final step of the ESCRT cascade involves the formation of a spiral polymer that lines the inside of the budding neck and constricts it (Step 3). Source: Modified from (Sundquist and Kräusslich, 2012)

The ESCRT complexes were originally identified in budding yeast, where they are important for multivesicular body (MVB) formation (Babst et al., 2002b, 2002a; Katzmann et al., 2001). ESCRT

enables endosomes containing ubiquitinylated transmembrane proteins to bud into MVBs before lysosomal degradation. However an increasing number of roles for ESCRT is being identified, including cytokinesis (Carlton and Martin-Serrano, 2007; Morita et al., 2007) and a range of other functions which have been recently reviewed (Campsteijn et al., 2016). Enveloped viruses hijack ESCRT because the events required for MVB formation and virus export are topologically equivalent, with the membrane budding away from the cytoplasmic face.

The ESCRT system is composed of four protein complexes (Table 1) which recruit each other in a sequential cascade. The membrane scission event is performed by ESCRT-III. ESCRT-III is composed of various charged multivesicular body proteins (CHMPs), the key components of which are CHMP2 and CHMP4, which form a spiral polymer to close the budding neck (Hanson et al., 2008). This is followed by recruitment of Vps4, an ATPase that recycles ESCRT-III components back into the cytoplasm (Lata et al., 2008).

The role played by ESCRT-II in HIV budding is controversial. Early studies showed that upon knockdown of ESCRT-II components virus release is not affected (Langelier et al., 2006; Pincetic et al., 2008). However, recent evidence in knockout cell lines has demonstrated that it may be required for optimal virus budding. A two-fold reduction in the virus release ratio was observed using siRNA knockdown of ESCRT-II and by CRISPR/Cas9 EAP45 knockout, however knockdown using shRNA left the release ratio unaffected despite reduced amounts of intracellular and extracellular Gag, which were attributed to an effect on transcription (Meng et al., 2015).

An *in vitro* study using purified proteins and giant unilamellar vesicles showed that recruitment of ESCRT-III to Gag complexes was dependent on the presence of ESCRT-II (Carlson and Hurley, 2012), however this is a simplified artificial system so interpretation of results must be performed with caution. Further work is needed to confirm the importance of ESCRT-II in HIV budding.

Table 1 – components of the ESCRT complexes in humans

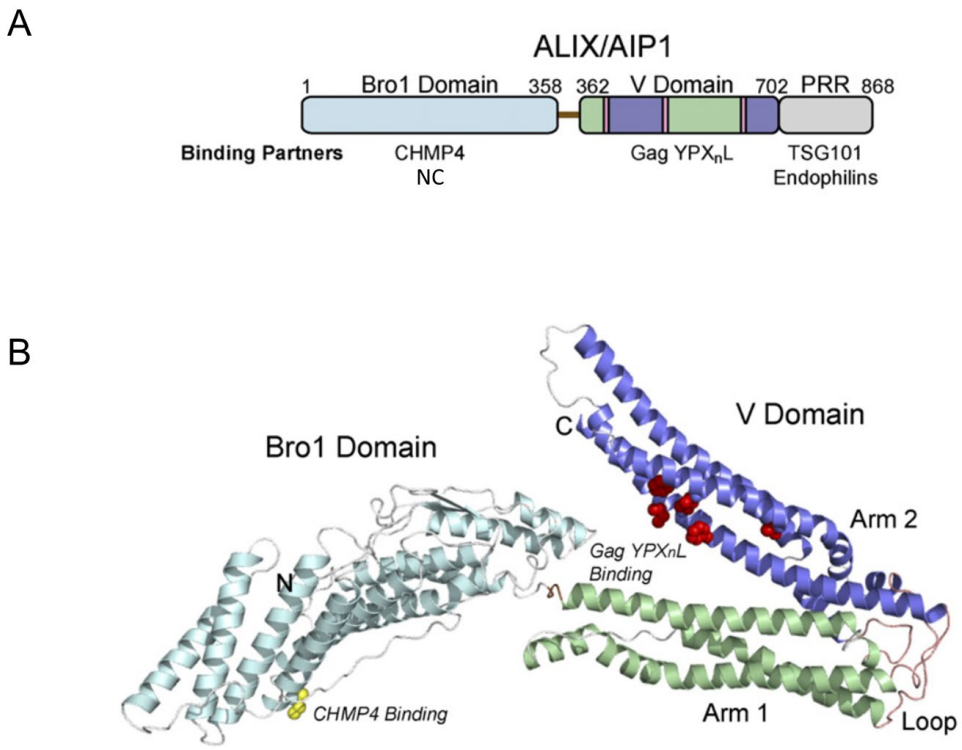
Complex name	Components
ESCRT-0	HGS; STAM1, 2
ESCRT-I	TSG101; VPS28; VPS37A, B, C, D; MVBB; UBAP1
ESCRT-II	EAP20; EAP30; EAP45
ESCRT-III	CHMP1A, B; CHMP2A, B; CHMP3; CHMP4A, B, C; CHMP5; CHMP6; CHMP7; IS1

ESCRT proteins are recruited through motifs in the p6 domain at the C-terminus of Gag (Figure 2). Deletion of p6 results in a severe reduction in particle release with virions remaining tethered to the plasma membrane (Göttlinger et al., 1991; Huang et al., 1995). The p6 domain is also called the “late domain” since mutation affects a late step in virus assembly.

The first motif to be identified as important for budding was the PTAP motif (Göttlinger et al., 1991; Huang et al., 1995; Martin-Serrano et al., 2001), which recruits tumour susceptibility gene 101 (TSG101), a component of ESCRT-I (Garrus et al., 2001; Martin-Serrano et al., 2001; VerPlank et al., 2001). In its prototypic role, TSG101 binds to ubiquitinylated transmembrane proteins and recruits downstream ESCRT complexes enabling endosomes to bud into MVBs (Katzmann et al., 2001). In the context of virus budding it is hijacked by Gag to recruit ESCRT to the plasma membrane (Garrus et al., 2001; VerPlank et al., 2001).

A second motif, YPX_nL, exists towards the C-terminus of p6 and recruits the host protein ALG-2 interacting protein X (ALIX) (Strack et al., 2003; Von Schwedler et al., 2003). ALIX is not a member of the ESCRT machinery but is termed an ESCRT-associated protein, and it also participates in MVB formation (Fujii et al., 2007). It has three major domains – the N-terminal Bro1 domain interacts with CHMP4 (Kim et al., 2005) and NC (Popov et al., 2008); the central V domain interacts with the YPX_nL motif in p6; and the proline-rich region interacts with TSG101 (Figure 10) (Fisher et al., 2007).

Figure 10 – Domains and crystal structure of ALIX



A) Domain organisation of ALIX with domain names labelled above and interaction partners labelled below. B) Crystal structure of ALIX. Colours of helices correspond with those regions indicated in A. Residues highlighted yellow and red are those involved in binding to CHMP4 (an ESCRT-III protein), and the YPXnL domain of Gag, respectively. A short hydrophobic linker connect the Bro1 domain with Arm 1 of the V domain.

Modified from (Fisher et al., 2007).

The ultimate outcome of both the TSG101 and ALIX pathways is to transiently recruit ESCRT-III components and Vps4 to the inside of the budding neck (Figure 11) (Prescher et al., 2015) at the end of Gag assembly (Jouvenet et al., 2011), enabling fission of budding necks and recycling of ESCRT components back into the cytoplasm.

The PTAP-TSG101 pathway is generally considered to be dominant, as disruption results in more severe defects in virus release in widely used cell lines (Fisher et al., 2007; Göttlinger et al., 1991; Huang et al., 1995). In 293T and HeLa cells, using full-length virus, disruption of the PTAP motif results in a 20-100-fold reduction in virus release (Fisher et al., 2007; Huang et al., 1995) whereas disruption of the YPX_nL motif results in a 2-3-fold reduction (293T) (Fisher et al., 2007), or in no reduction (HeLa) (Demirov et al., 2002).

However, in a minimal Gag construct lacking the globular domain of MA and the C-terminus of CA, YPX_nL mutation results in a 10-fold reduction in release from HeLa cells, suggesting that other factors compensate for YPX_nL mutation in the context of a full-length virus (Strack et al., 2003).

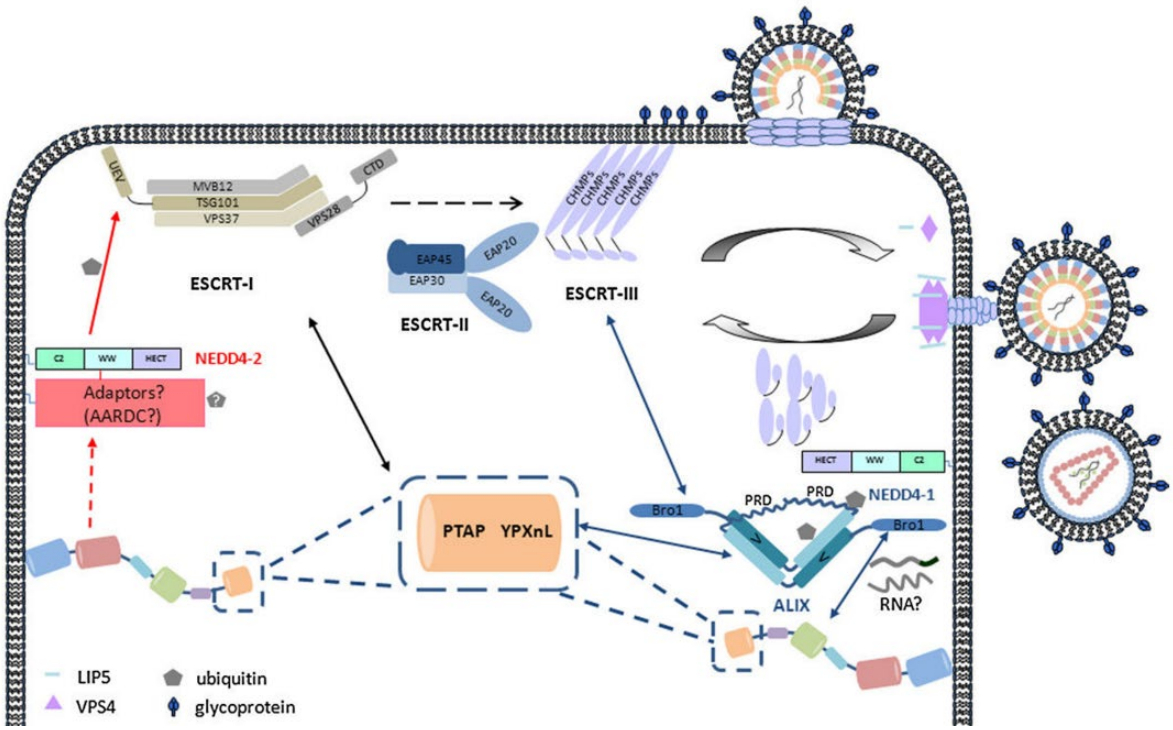
Whilst being useful investigational tools, 293T and HeLa cells may not accurately replicate the physiological conditions encountered by the virus in natural infection. Fujii et al introduced mutations into the YPXnL motif and evaluated the effect on virus release in a HeLa cells in addition to physiologically relevant primary T cells and macrophages. As observed before and by others, in HeLa cells ALIX binding site mutations had no effect on virus release (except for a 5-fold reduction in the Y36A mutant), whilst $PTAP^-$ mutation resulted in a 10-fold reduction. In the Jurkat T-cell line and in monocyte-derived macrophages (MDMs), the Y36A mutation resulted in greater (Jurkat) or equivalent (MDMs) release defects than the $PTAP^-$ mutant.

In a virus replication assay, ALIX binding site mutations resulted in more severe defects in replication kinetics than $PTAP^-$ mutation, whilst the reverse was true in MDMs. The authors hypothesised that the relative importance of the TSG101 and ALIX pathways may be due to differences in expression levels between cell types, however the Jurkat cell line expresses higher levels of TSG101 but appears to be more dependent on the ALIX pathway (Fujii et al., 2009). The block in virus release caused by interference with the PTAP-TSG101 pathway can be rescued by overexpressing ALIX (Fisher et al., 2007; Usami et al., 2007), providing that the YPX_nL motif and CHMP4 binding site are intact. Therefore, ALIX can compensate for a defective TSG101 pathway by providing an alternative link to ESCRT-III. ALIX therefore appears to play a vital role in circumstances of low TSG101 availability or disruption of Gag-TSG101 binding (Usami et al., 2007).

Microscopy studies initially showed that GFP-tagged ALIX is recruited progressively to Gag assembly sites (Jouvenet et al., 2011). A later study found that the GFP tag disrupts ALIX function, but that introduction of a linker corrects this defect. These wild type-like ALIX molecules were recruited to the

budding neck transiently at the end of assembly (Ku et al., 2014). In contrast TSG101 appears to accumulate progressively with Gag assembly (Bleck et al., 2014). ALIX and TSG101, despite binding in very similar positions in p6, are positioned distinctively in assembling virions – ALIX is positioned more peripherally than TSG101, which is more central (Bleck et al., 2014).

Figure 11 – Role of ESCRT system in HIV-1 release



The PTAP and YPXnL motifs in the p6 domain of Gag recruit TSG101, a member of ESCRT-I, and ALIX, an ESCRT-associated protein, respectively. Both pathways converge on recruitment of ESCRT-III, components of which forms spiral filaments that constrict and sever the budding neck, before recycling back into the cytoplasm by VPS4.

Source: (Meng and Lever, 2013).

The practice of usurping the ESCRT pathway by hijacking TSG101 and ALIX is not limited to HIV or even to retroviruses. ESCRT is the only well characterised cellular pathway which facilitates budding away from the cytoplasmic face, and as such it is hijacked by a plethora of viruses. The TSG101-binding P(T/S)AP motif first identified in HIV-1 can be found in the structural proteins of filoviruses, arenaviruses, rhabdoviruses and reoviruses.

The ALIX-binding YPXL motif first identified in Equine Infectious Anemia Virus has since been found in paramyxoviruses, arenaviruses, flaviviruses, hepadnaviruses, herpesviruses and tombusviruses (Votteler and Sundquist, 2013). ALIX is also important for Ebola virus release – the VP40 matrix protein of the virus possesses a YPXL/I consensus motif that interacts with ALIX. siRNA knockdown of ALIX

reduced virus release, and overexpression of ALIX rescued production of the budding deficient PTAP/PPEY L-domain deletion mutant (Han et al., 2015).

1.15. Virus maturation

During and after budding, PR cleaves Gag and GagPol at highly conserved recognition sites (Figure 12) (Kaplan et al., 1994; Lee et al., 2012; Mattei et al., 2016), each with a consensus shape recognised by PR despite differences in sequence (Prabu-Jeyabalan et al., 2002). Cleavage at each site occurs at a different rate (Pettit et al., 2005) so processing takes place in a particular order (Figure 12) which if disturbed by cleavage site mutations or pharmacological agents is extremely detrimental to virus infectivity (Li et al., 2003; Nowicka-Sans et al., 2016; Wiegers et al., 1998). Gag is cleaved into MA, CA, SP1, NC, SP2 and p6 (Figure 12), whilst Pol is cleaved into PR, RT and IN.

Gag processing triggers a structural rearrangement of the viral core termed maturation, which is essential for the virus to become infectious (Göttlinger et al., 1989; Kohl et al., 1988; Peng et al., 1989). The electron dense outer shell is replaced by a conical inner core (Figure 4 and Figure 12). Gag processing also triggers dimer maturation (Figure 7) because the mature NC protein acts as a chaperone facilitating formation of the extended dimer (Darlix et al., 1990; Muriaux et al., 1996).

Figure 12 – Gag processing and virion maturation

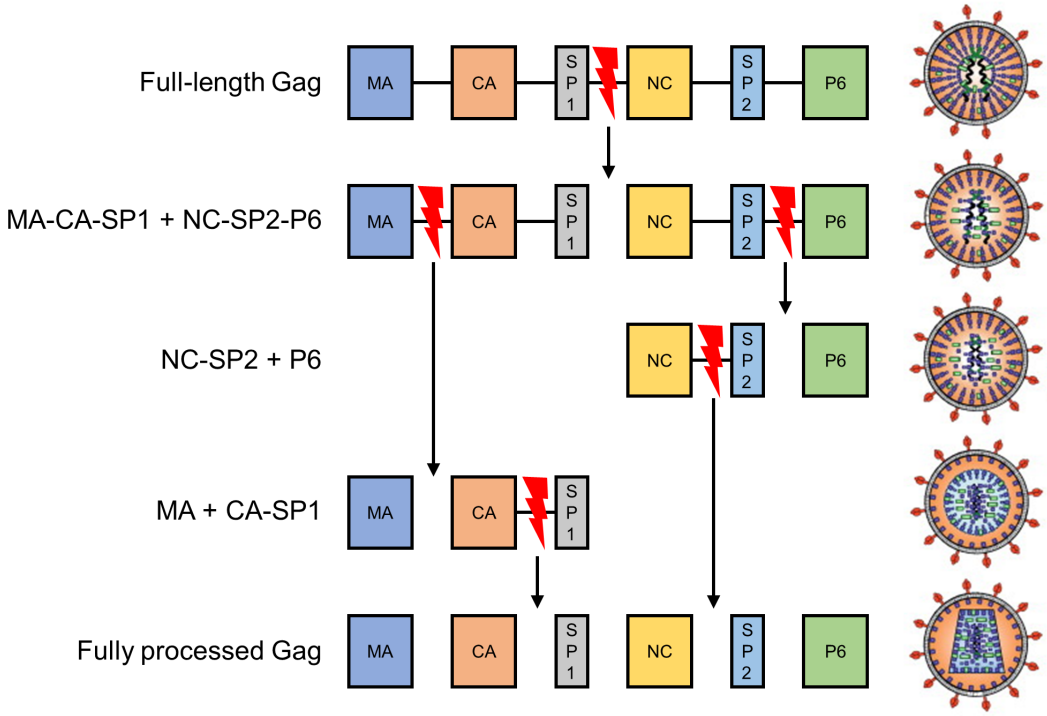


Illustration of the stages of Gag processing and the accompanying changes in virion morphology. Red lightning strikes represent cleavage events, which appear on the page in the order they occur in the virion. This ordered processing occurs because the 5 cleavage sites between the 6 domains of Gag are cleaved by PR at different rates. Gag processing causes the viral core to undergo structural rearrangement, with the final CA-SP1 cleavage being particularly important for the formation of a conical core. To aid interpretation of the figure the domains of Gag are not drawn to scale.

The PR enzyme is a homodimer, with each monomer contributing an aspartic acid residue critical for formation of the active site (Wlodawer and Gustchina, 2000). As PR is initially embedded as a domain within GagPol this necessitates the formation of a GagPol dimer to commence Gag processing. The PR domains in the GagPol dimer self-excise through an intramolecular cleavage (Pettit et al., 2004) to become a much more stable molecule with greater enzymatic activity (Lee et al., 2012).

The mechanism underpinning the non-infectious nature of immature virions remained unknown for some time, but several investigators have found that blocking Gag processing prevents virions from fusing with target cells (Chojnacki et al., 2012; Murakami et al., 2004; Wyma et al., 2004). Two mechanisms have been described that explain this phenomenon, although the two are not mutually exclusive and are likely to both be required for efficient (Chojnacki et al., 2017). Studies by the Rousso group found that maturation of the core reduced virion stiffness and increased fusogenicity, likely due to the dramatic structural reorganisation of Gag that takes place. Curiously the increased stiffness of immature particles was abolished by removal of the cytoplasmic tail of Env, suggesting that Env may

help to stabilise the immature Gag shell in addition to its primary role of facilitating fusion. The authors proposed that this stiffness switch acts to prevent the entry of aberrantly formed virions (Kol et al., 2007; Pang et al., 2013).

Chojnacki *et al* also observed a role for the cytoplasmic tail of Env in fusion, however in their study they were using stimulated emission depletion microscopy to study the effect of maturation on the localisation of Env trimers on the viral membrane. In mature particles Env trimers clustered together into a single focus, whereas in immature particles several foci were observed. They found that proteolytic processing of Gag, particularly at MA-CA, triggered clustering of the limited number (7–14) of Env trimers on the virion surface. Deletion of the cytoplasmic tail of Env rescued the clustering defect in immature virions. They proposed a model whereby the cytoplasmic tail of Env interacting with the immature Gag lattice traps the trimers on the virion surface. Upon maturation Env molecules are liberated from this trapped state to form a single focus enabling efficient interaction with CD4 (Chojnacki et al., 2012).

CA-SP1 processing is a particularly important step as it completes the dramatic structural rearrangement of the viral core (Figure 12) (Wiegers et al., 1998). CA-SP1 cleavage sites are not accessible to PR in the immature virion as they are folded into α-helices forming 6-helix bundles which support the hexameric lattice. Only upon unwinding of the helix is the cleavage site exposed, making CA-SP1 cleavage the slowest of all the processing steps (Bayro et al., 2016; Schur et al., 2016; Wagner et al., 2016). CA-SP1 processing can be specifically inhibited by a new experimental class of ARVs, termed the maturation inhibitors (MIs), although none have been authorised for clinical use yet. The prototypic MI is PA-457, otherwise known as bevirimat (Li et al., 2003).

Gag processing and budding are tightly linked events, although the mechanism linking virus budding and Gag processing remains a mystery. This is mainly due to the fact they occur simultaneously and interfering with one process often interferes with the other - processing is disturbed by disruption of late domain function (Fisher et al., 2007; Huang et al., 1995) and budding is inhibited when PR is overexpressed (Karacostas et al., 1993) or inhibited (Kaplan et al., 1994).

PTAP and YPX_nL mutants exhibit Gag processing defects characterised by accumulation of the MA-CA-SP1 and CA-SP1 processing intermediates (Fisher et al., 2007; Huang et al., 1995). Interference further down the ESCRT pathway, such as knockdown of ESCRT-II (Meng et al., 2015), ESCRT-III (Morita et al., 2011) or over-expression of dominant-negative VPS4 protein (Garrus et al., 2001), also causes defective CA-SP1 processing and reduced virus release. As disruption of all stages in the pathway causes reduced CA-SP1 processing, this suggests that this crucial cleavage event is triggered during the final step of the ESCRT cascade.

Inhibition of PR rescues the budding defect of late domain mutants (Huang et al., 1995) by preventing over-processing of Gag caused by delayed release (Ott et al., 2009). On the other hand, introduction of PR into late domain mutant Gag-expressing cells inhibits release because Gag and GagPol are processed before closure of the budding neck, leading to diffusion of critical virion components back into the cytoplasm (Bendjennat and Saffarian, 2016).

This highlights a race that takes place between closure of the budding neck and activation of PR. Interventions that disrupt the fragile balance between budding and processing result in severe attenuation of virus replication.

1.16. Replication cycle summary

In summary, the wealth of biochemical data and recent data from microscopy and immunoprecipitation studies suggests that the following sequence of events occurs during the late stages of virus assembly:

1. Following export of gRNA from the nucleus, dimerisation occurs between complementary DIS sequences stimulated by Gag molecules acting as chaperones. The NC domain of Gag is bound to specific sequences (psi, the RRE and the 3'UTR).
2. This results in the formation of Gag-gRNA complexes in the cytoplasm; oligomers of Gag bound to gRNA, which traffic to the plasma membrane by diffusion.
3. Upon arrival at the plasma membrane Gag anchors into the membrane through the MA domain, stabilising the complex in the membrane. Additional dimerisation can occur on the plasma membrane. Gag's binding preference switches from specific motifs to more general binding throughout the genome.
4. Recruitment of monomeric Gag from the cytoplasm results in multimerisation to form an immature virion particle. Multimerisation is mediated by contacts between CA and NC although the RNA interactions required for NC-NC interactions are non-specific.
5. Following ESCRT recruitment the virus is released from the cell and maturation occurs. The dramatic structural changes are accompanied by changes in the binding specificity of Gag to gRNA, which again binds to specific regions (psi, RRE, 3'UTR) possibly to prepare the genome for reverse transcription to begin.

1.17. Observations and findings which led to this study

The experiments performed herein were based on the observation that mutation of SL1 in the 5'UTR results in aberrant Gag processing and formation of immature virions (Liang et al., 1999a). An understanding of the mechanism linking an RNA structure in psi with Gag processing could potentially lead to new insights into the process of virus assembly.

Additionally, despite the efficacy of ARVs, the genetic flexibility of the virus means that the evolution of resistance mutations is a serious problem. RNA structures in the 5'UTR are some of the most highly conserved regions of the virus genome, so an improved understanding of the functions of these regions may provide insight into new mechanisms which can be targeted therapeutically (Le Grice, 2015).

2. Materials and methods

2.1. Solutions

All solutions were prepared with MilliQ water (Merck Millipore). Recipes are provided in Table 10 in Appendix A.

2.2. Plasmid constructs and cloning

2.2.1. Plasmid DNA constructs obtained prior to this study

All plasmids express ampicillin resistance genes.

pSVC21.BH10-WT, an infectious molecular clone of HIV-1 group M subtype B (Fisher et al., 1985; Ratner et al., 1985) and the DIS, Stem B and Loop B mutants were gifts from Dr. Michael Laughrea (McGill University, Canada). All plasmid nucleotide positions mentioned hereafter refer to the reference sequence (accession number K03455.1). The plasmids had been made suitable for containment level 2 work by Dr. Bo Meng by deleting a part of the coding region in *Env* (7042-7621) using the restriction enzyme *BglII*. The DIS mutant has the C residue at position 712 of the reference sequence substituted for a G, the Stem B mutant has nucleotides 697-700 deleted and the Loop B mutant has nucleotides 725-727 deleted. These mutations affect the local structure of SL1 in the 5'UTR of the transcribed gRNA. These plasmids are termed pSVC21.BH10-*Bgl*/(mutant name).

pCMV-VSV-G, a gift from Dr. Laurence Tiley, encodes the vesicular stomatitis virus G glycoprotein (VSV-G) under the control of a CMV promoter.

pBlueScript (Short et al., 1988) is a cloning vector. It was used in this study as an empty vector for equilibrating the amount of DNA used between different transfection conditions.

pGEM-T Easy (Promega) is approximately a 3-kilobase (kb) linearised cloning vector. It has a 3' terminal thymidine at each end, enabling ligation with PCR products generated by *Taq* polymerases and reducing the risk of vector recircularisation.

pBJ5-ALIX-HA was a gift from Professor Heinrich Gottlinger (University of Massachusetts Medical School, USA) (Strack et al., 2003). It encodes full-length ALIX with an HA tag at the C-terminus, cloned into the mammalian expression vector pBJ5. Expression is under control of an SRα promoter (Takebe et al., 1988).

pSVC21.BH10-*Bgl*/PTAP$^-$ was constructed by Dr. Bo Meng by mutating in the region of nucleotides 2152-2163 (CCA ACA GCC CCA → CTA ATT CGC CTA) as described in (Huang et al., 1995). The proline-threonine-alanine-proline (PTAP) motif in the p6 domain of Gag is changed to leucine-isoleucine-arginine-leucine (LIRL). This prevents TSG101 binding whilst maintaining the integrity of the amino acid sequence of the overlapping *Pol* reading frame (Huang et al., 1995).

pSVC21.BH10-*Bgl*/YPX$_n$L$^-$ was constructed by Dr. Bo Meng by mutating in the region of nucleotides 2239-2244 (TAT CCT → TTA GTT) as described in (Dussupt et al., 2009). The tyrosine-proline-X$_n$-leucine (YPX$_n$L – where X is any amino acid and n = 1, 2 or 3 (in the case of BH10 X$_n$ = leucine-threonine-serine)) motif in p6 is altered to Leucine-Valine-X$_n$-Leucine (LVX$_n$L), preventing ALIX binding whilst again maintaining the *Pol* amino acid sequence (Dussupt et al., 2009).

pSVC21.BH10-*Bgl*/PTAP$^-$/YPX$_n$L$^-$, containing both above-mentioned sets of mutations, was constructed by Dr. Bo Meng by mutating the YPX$_n$L domain in pSVC21.BH10-*Bgl*/PTAP$^-$.

pSVC21.BH10-*Bgl*/F16A contains a substitution of the 16th amino acid in NC from phenylalanine to alanine and was constructed by Dr. Bo Meng.

pJHIV-1-Luc and pJHIV-1 contain nucleotides 1-336 of HIV-1 HXB2 upstream of a sequence encoding firefly luciferase (pJHIV-1-Luc) or nucleotides 47-889 of influenza A virus segment 8 (pJHIV-1). pJHIV-1-ΔΨ construct is identical to pJHIV-1 except that nucleotides 224-336, which form SLs 1-4, are deleted. Production of the pJHIV-1 and pJHIV-1-ΔΨ constructs has been previously described (Anderson and Lever, 2006).

2.2.2. Plasmid DNA constructs produced during this study

The following plasmids were constructed using the cloning methods described in the rest of this section.

pSVC21.BH10-*Bgl*/LD4 has nucleotides 718-731 deleted. It was produced by PCR site-directed mutagenesis (see Section 2.2.7).

pSVC21.BH10-*Bgl*/MP2/MNC/(WT, DIS, Stem B, Loop B, LD4, PTAP$^-$, YPX$_n$L$^-$ or PTAP$^-$/YPX$_n$L$^-$) contain two mutations in Gag. The MNC mutation substitutes a C for T at position 1989 changing the amino acid encoded from threonine to isoleucine in the NC domain of Gag. The MP2 mutation substitutes C for T at position 1913, also resulting in a change of amino acid from threonine to isoleucine in the SP1 (also known as p2) domain of Gag. The mutations were introduced into pGEM-T-UTR-WT subclone (see Section 2.2.8) before transfer back to the respective pSVC21.BH10-*Bgl* provirus constructs (WT, SL1 mutants and PTAP$^-$/YPX$_n$L$^-$ mutants) (see Section 2.2.9).

pSVC21.BH10-*Bgl*/LD4/PTAP$^-$ contains the LD4 deletion (718-731) in addition to the PTAP- mutation (PTAP→LIRL). It was produced by digesting pSVC21.BH10-*Bgl*/PTAP$^-$ and pSVC21.BH10-*Bgl*/LD4 with *Apa*I and *Xcm*I (see Section 2.2.5) and ligating the PTAP- insert into the LD4 vector (see Section 2.2.6).

pSVC21.BH10-*Bgl*/AAG contains a mutation in the Gag start codon, which is mutated from ATG to AAG (T→A at position 791). pSVC21.BH10-*Bgl*/AAG/CA contains the start codon mutation as well as a second mutation in the CA domain of Gag. The mutation is an insertion of 4 nucleotides (CTAG) between nucleotides 1511 and 1512. This introduces a stop codon and a frameshift preventing

translation from alternative start sites. The mutations were introduced into pGEM-T-UTR-WT (see Section 2.2.8) before transfer back to the pSVC21.BH10-*Bgl* provirus construct (see Section 2.2.9).

2.2.3. Agarose gel electrophoresis and extraction of DNA

To assess the size and concentration of plasmids, PCR products and restriction digest products, samples were mixed with GelPilot 5X DNA Loading Dye (Qiagen) and electrophoresed in agarose gels. Gels were made by dissolving 0.5 g of agarose (Sigma Aldrich) in 50 ml of tris-borate-EDTA (TBE) buffer by boiling, allowing to cool slightly and adding 5 µl of SYBR Safe DNA Gel Stain (Thermo Fisher), and then pouring into a gel tank to set. When gels had set, they were immersed in 0.5X TBE running buffer and samples were loaded, alongside 5 µl of HyperLadder 1kb (Bioline), before running at 60 V for the time required to separate the relevant bands. Gels were visualised, and images captured in a gel imaging box.

DNA was purified when required by placing the gel on a UV transilluminator, cutting out the relevant bands using a scalpel and processing the gel slices using the QIAquick Gel Extraction Kit (Qiagen), following the manufacturer's protocol.

2.2.4. Polymerase chain reaction

Polymerase chain reaction (PCR) mixes consisted of 1 µl of 10 ng/µl template DNA, 2 µl of 10 µM forward and 2 µl of 10 µM reverse primers (**Table 2**), 1 µl of deoxyribonucleotide triphosphate (dNTP) mix and 5 µl of 5X GoTaq buffer (Promega) with 0.25 µl of GoTaq G2 (Promega) or, for high fidelity reactions, 2.5 µl of 10X *Pfu* buffer (Promega) with 0.5 µl of *Pfu* (Promega), made up to 25 µl with nuclease-free water.

Thermal cycling conditions used (unless otherwise specified) were: initial denaturation at 95 ^{o}C for 2 minutes, followed by 35 cycles of denaturation at 95 ^{o}C for 30 seconds, annealing at 52 ^{o}C for 30 seconds and extension at 73 ^{o}C for 2 minutes, followed by a final extension at 73 ^{o}C for 5 minutes.

To purify PCR product from the reaction mixture, the QIAquick PCR Purification Kit (Qiagen) was used, following the manufacturer's protocol. Samples were taken and run on agarose gels to verify their size and purity, and the purified PCR products were stored at -20 ^{o}C.

Table 2 – primers used for PCR

PCR primer name	Sequence (5'-3')
581F	CTGGTAACTAGAGATCCCTC
2054R	TGGTGTCCTTCCTTTCCACATTTC
1878F	GGCTGAAGCAATGAGCCAAG
2984R-T7	TAATACGACTCACTATAGGGAGACTGATATCTAATCCCTGGTG

2.2.5. Restriction digestion

Restriction digest reactions were composed of 1 µg of DNA, 2 µl of 10X MULTI-CORE buffer (Promega) or 10X CutSmart buffer (New England Biolabs) and 10 units (U) of each restriction enzyme, made up to 20 µl with nuclease-free water. 1 µl (10 U) of thermosensitive alkaline phosphatase (TSAP) (Promega)

was also added to the mixture if the digest product was to be subsequently used for ligation, to prevent recircularisation. Reactions were incubated for 3 hours at the required temperature (Table 3). To inactivate enzymes the reaction mixtures were incubated at 75 ^{o}C for 15 minutes.

Digest products were analysed using agarose gel electrophoresis and purified if required. Digestion with multiple enzymes was used as a diagnostic tool to check that newly produced plasmids had the same digest pattern as the parental plasmids from which they were derived.

Table 3 – restriction enzymes used

Enzyme (manufacturer)	Buffer (manufacturer)	Operating temperature (^{o}C)
*Apa*I (Roche)	MULTI-CORE (Promega)/ CutSmart (New England Biolabs)	37
*BssH*II (New England Biolabs)	MULTI-CORE	50
*Xcm*I (New England Biolabs)	CutSmart	37
*Spe*I (New England Biolabs)	MULTI-CORE	37

2.2.6. Ligation

Ligation reactions were set up with a 3:1 ratio of insert:vector. Reactions were made up of 2 µl of 10X Ligase buffer (Promega), 50 ng of vector, the required amount of insert and 0.5 µl of T4 DNA Ligase (Promega) made up to 20 µl with nuclease-free water. Control reactions were also set up which lacked insert, so that following bacterial transformation the number of colonies could be compared to give an indication of whether the ligation was successful or not. Ligation and control reactions were incubated for 3 hours at room temperature or overnight at 4 ^{o}C and then inactivated by incubating for 10 minutes at 70 ^{o}C.

2.2.7. PCR site-directed mutagenesis

pSVC21.BH10-*Bgl*/LD4 was produced using PCR site-directed mutagenesis. The primers DIS-R and Apa-A (Liang et al., 1998) contain restriction sites for *BssH*II and *Apa*I respectively (Table 4), allowing the PCR product to be restriction digested and ligated into the vector. The DIS-R primer contains the LD4 deletion (nucleotides 718-731) to be introduced into PCR products. The PCR reaction was set up using *Pfu* to minimise the risk of unwanted mutations being introduced, and modifying the annealing temperature to 65 ^{o}C and the extension time to 75 seconds.

Table 4 – primers used for PCR site-directed mutagenesis

Mutagenic primer name	Sequence (5'-3') (* = restriction enzyme cut site)
DIS-R	GCTGAAG*CGCGCAGGGCGGCGACTGGTGAGTACGCC
Apa-A	CCTAGGGGCC*CTGCAATTTCTG

The purified PCR product and pSVC21.BH10-*Bgl* were then digested with *Apa*I and *BssH*II in a sequential digest due to the different operating temperatures of the two enzymes. First an *Apa*I digest was performed, followed by inactivation, then *BssH*II was added. The digests were run on agarose gels and

purified. Ligation reactions were set up followed by transformation of 5 µl of the ligation mixture into bacteria (see Section 2.3.1).

The sizes of cultured plasmids were checked by restriction digestion and agarose gel electrophoresis before verifying the introduction of the mutation by sequencing (GATC Biotech).

2.2.8. Production of pGEM-T-UTR subclone

A pGEM-T Easy vector-based subclone containing nucleotides 581-2054 was produced to provide a construct which could be used for mutagenesis and transfer back into the pSVC21.BH10-*Bgl* plasmids. Primers 581F and 2054R (Table 2) were used to amplify this region from pSVC21.BH10-*Bgl* by PCR, followed by purification of the product before ligation of into the pGEM-T Easy vector using the manufacturer's protocol. Following transformation and culture, the product size (4.5 kbp) was confirmed by agarose gel electrophoresis and sent for sequencing with primer 581F. The plasmid was termed 'pGEM-T-UTR-WT'.

2.2.9. Subclone mutagenesis and transfer back to proviral vector

To produce the other mutants in this study (MP2/MNC, AAG and AAG/CA mutations), whole plasmid site-directed mutagenesis was performed on the subclone pGEM-T-UTR-WT. The primers pairs used both contain the desired mutation and are complementary to each other. Unlike PCR, the reaction is not an amplification since only the parental plasmid can be used as a template (nicks created in newly formed plasmids preclude their use as a template for further amplification).

Reaction mixtures were composed of 5 µl of 10X *Pfu* buffer, 1 µl each of 10 µM NC-F and NC-R primers, 1 µl of dNTP mix, 5 µl of 10 ng/µl of pGEM-T-UTR subclone and 0.5 µl of *Pfu* polymerase made up to 50 µl with nuclease-free water.

Thermal cycling conditions used were: initial denaturation at 95 ^{o}C for 1 minute, followed by 12 cycles of denaturation at 95 ^{o}C for 30 seconds, annealing at 65 ^{o}C for 1 minute and extension at 68 ^{o}C for 8 minutes. Following this 1 µl (20 U) of *Dpn*I (New England Biolabs) was added to the reaction and incubated for 1 hour at 37 ^{o}C to digest the wild type parental methylated plasmid.

Plasmids were then transformed, cultured and sent for sequencing to confirm the introduction of the desired mutations.

To introduce the MNC mutation into the subclone, primers NC-F and NC-R were used, and to introduce the MP2 mutation, primers P2-F1 and P2-R1 were used. This plasmid was termed 'pGEM-T-UTR-MP2/MNC'.

To produce a construct which does not express Gag, site-directed mutagenesis was first performed on the subclone pGEM-T-UTR-WT using primers Gag AUGmut F and Gag AUGmut R (Table 5), changing the Gag start codon nucleotide sequence from ATG to AAG. An additional mutation was made using

primers CA Stop F1 and CA Stop R1 (Table 5), which introduced a premature stop codon and a frameshift in the CA domain of Gag. This plasmid was termed 'pGEM-T-UTR-AAG/CA'.

To transfer mutated subclone sequences back into the pSVC21.BH10-*Bgl* provirus construct, primers 518F and 2054R (Table 2) were used in a PCR reaction with *Pfu*, followed by purification. Both the PCR product and the pSVC21.BH10-*Bgl* plasmid were then digested with *Apa*I and *Spe*I, with TSAP present to prevent recircularisation of the vector during ligation. The digest was run on an agarose gel and the largest restriction fragment from the insert digest (1524 nucleotides) was gel extracted and purified. This was then ligated back into pSVC21.BH10-*Bgl* followed by transformation, culture and sequencing with primers 518F, 1616F or 2054R. Plasmid size and integrity was assessed by digesting with various restriction enzymes alongside pSVC21.BH10-*Bgl* plasmid to ensure that the same digest pattern was obtained.

Table 5 – primers used for subclone mutagenesis

Mutagenic primer name	Sequence (5'-3')
NC-F	GGCACATAGCCAGAAATTGCAG
NC-R	CTGCAATTTCTGGCTATGTGCC
P2-F1	GCCAAGTAACAAATTCAGCTATCATAATGATGCAGAGAGGC
P2-R1	GCCTCTCTGCATCATTATGATAGCTGAATTTGTTACTTGGC
Gag AUG mut F	GGAGGCTAGAAGGAGAGAGAAGGGTGCGAGAGCG
Gag AUG mut R	CGCTCTCGCACCCTTCTCTCTCCTTCTAGCCTCC
CA stop F1	GGAAGTGACATAGCAGGAACTACTAGCTAGTACCCTTCAGG
CA stop R1	CCTGAAGGGTACTAGCTAGTAGTTCCTGCTATGTCACTTCC

2.3. Bacterial expression of plasmids and sequencing

2.3.1. Bacterial transformation and culture

For routine plasmid transformation, the laboratory stock of the DH5α strain of *E. coli* was used (Taylor et al., 1993). 5 µl of 10 ng/µl plasmid was added to a 100 µl aliquot of bacterial cells, which had been stored at -80 ^{o}C and allowed to thaw on ice for 10 minutes. The mixture was gently flicked to mix and then left on ice for a further 30 minutes. Bacteria were then heat-shocked by incubating at 42 ^{o}C for 40 seconds, and on ice for 2 minutes, before adding 1 ml of Luria Bertani (LB) and incubating in a shaking incubator at 37 ^{o}C and 220 rpm for 1 hour.

For transformation of ligation mixtures XL10-Gold ultracompetent cells (Stratgene) were used. Cells which had been stored at -80 ^{o}C were thawed on ice for 10 minutes. 50 µl aliquots were added to pre-chilled round-bottom polypropylene tubes on ice, and 2 µl of 2-mercaptoethanol was added to each aliquot before swirling gently and incubating on ice for 10 minutes, swirling every 2 minutes. 5 µl of ligation mixture were added to each aliquot before swirling and incubating on ice for 30 minutes. Bacteria were then heat-shocked by incubating at 42 ^{o}C for 30 seconds, and on ice for 2 minutes, before plating directly.

Transformation mixtures were then plated onto LB agar supplemented with 100 µg/ml ampicillin and incubated overnight at 37 ^{o}C. The following day colonies were counted on the ligation plate and the control plate, where ligation was performed without addition of insert. If the number of colonies on the ligation plate was greater than the control plate, this indicated that the ligation had been successful.

Colonies were picked using a sterile pipette tip and inoculated into 5 ml aliquots of LB medium supplemented with 100 µg/ml ampicillin for overnight culture in a shaking incubator at 220 rpm and 37 ^{o}C. If a larger quantity of plasmid DNA was required 1 ml of 5-ml culture was used to inoculate 400 ml of LB medium supplemented with 100 µg/ml ampicillin.

2.3.2. Purification of plasmids from culture

Plasmid DNA was purified from 5 ml and 400 ml overnight cultures using the QIAprep Spin Miniprep Kit (Qiagen) and the Qiagen Plasmid Maxi Kit (Qiagen) respectively, following the manufacturer's protocol. To determine the concentration of purified DNA a spectrophotometer (NanoDrop) was used.

2.3.3. Producing and reviving glycerol stocks

For long-term storage, glycerol stocks of transformed bacteria were made by mixing 500 µl of bacterial culture with 500 µl of 50% glycerol and storing at -80 ^{o}C. Bacteria were later recovered by stabbing into the frozen stock with a sterile pipette tip and inoculating into 5 ml of LB medium supplemented with 100 µg/ml ampicillin for overnight culture. DNA was then purified as described above.

2.3.4. Plasmid DNA sequencing

Custom sequencing primers were obtained from Sigma Aldrich as dried pellets, which were resuspended to 100 µM with nuclease free water and stored at -20 ^{o}C.

For sequencing reactions plasmids were diluted to 100 ng/µl and sequencing primers to 10 µM in a volume of 20 µl. Samples were sent to GATC Biotech for Sanger sequencing. Sequences were returned in FASTA format and analysed using UGENE (Unipro) software.

Table 6 – primers used for sequencing

Sequencing primer name	Sequence (5'-3')
1616F	TGTATAGCCCTACCAGCATTCT
581F	CTGGTAACTAGAGATCCCTC
2054R	TGGTGTCCTTCCTTTCCACATTTC

2.4. Cell culture and transfection

2.4.1. Maintenance of cell lines

293T cells (ATCC) and TZM-bl cells (NIH AIDS Reagent Program) were maintained in Dulbecco's modified Eagle medium (DMEM) (Thermo Fisher). The medium was supplemented with 10% (v/v) fetal bovine serum (FBS) (Thermo Fisher).

Cells were incubated in T75 flasks in 5% CO_2 at 37 ^{o}C, and were passaged 1:10 every 3-4 days. To passage cells the medium was discarded and the cells were washed gently with phosphate-buffered saline (PBS) before adding 1 ml of trypsin (Thermo Fisher), and incubating for 5 minutes at 37 ^{o}C. Trypsinised cells were then resuspended to a volume of 10 ml with fresh medium. 1 ml of resuspended cells were transferred to a new T75 flask, and 12 ml of fresh medium was added. Cells were discarded and fresh stocks resuscitated before cells reached passage number 30.

2.4.2. Freezing and resuscitation of cells

To freeze down cells for long term storage in liquid nitrogen, 4 ml of cultured cell suspension and 0.5 ml of FBS were mixed, before adding 0.5 ml of dimethyl sulfoxide (DMSO) dropwise. 1 ml aliquots were added to cryovials and cooled at the recommended rate of 1 ^{o}C/min using a Mr Frosty freezing container (Thermo Fisher) following the manufacturer's protocol. After overnight freezing the cryovials were transferred to liquid nitrogen for storage.

To thaw frozen cells, cryovials were removed from liquid nitrogen and warmed in a 37 ^{o}C water bath with gentle shaking. The 1 ml aliquots were transferred to 15 ml conical centrifuge tubes and 5 ml of culture medium was added before centrifuging at 1000 rpm for 2 minutes in a benchtop centrifuge. The supernatant was carefully removed before resuspending the cell pellet in 5 ml of culture medium and transferring to a T25 tissue culture flask and incubating as described above.

2.4.3. DNA and siRNA transfection overview

24 hours prior to transfection, 293T cells were seeded at a confluence of 40% in 6-well (≈ $2.4x10^6$ cells per well) or 24-well plates (≈ $4.8x10^5$ cells per well) to reach a confluence of 80% at transfection. For plasmid transfection DNA was mixed with *Trans*IT-LT1 Transfection Reagent (Mirus Bio) (see Appendix B for a comparison of different transfection reagents) at a ratio of 3 µl of transfection reagent per 1 µg of DNA. 1µg of DNA and 3 µl of transfection reagent were used per well of a 6-well plate (unless otherwise specified), and 200 ng of DNA and 0.6 µl of transfection reagent were used per well of a 24-well plate. Transfection reagent and DNA were mixed in a volume of 250 µl (for 6-well plates) or 50 µl (for 24-well plates) of Opti-MEM Reduced Serum Medium (Thermo Fisher). Mixtures were incubated at room temperature for 15-30 minutes. The transfection mixture was then added dropwise to the culture media, before returning the cells to the incubator.

For siRNA transfection 2µl of 20 µM Stealth RNAi siRNA (Thermo Fisher) was mixed with 1 µl of *Trans*IT-TKO Transfection Reagent (Mirus Bio) and 50 µl Opti-MEM per well of a 24-well plate, before incubating and adding to culture media as described above. 24 hours' post-transfection the siRNA transfection was repeated.

2.4.4. *In trans* expression of gRNA/5'UTR

293T cells seeded in 24-well plates were transfected with 100 ng of pSVC21.BH10-*Bgl* proviral DNA, and 100 ng of pSVC21.BH10-*Bgl*/AAG/CA, pJHIV-1, pHIV-1-Luc, pJHIV-1-ΔΨ or pBlueScript.

2.4.5. Knockdown of ALIX and TSG101 in the presence and absence of MP2/MNC compensatory mutations

293T cells seeded in 24-well plates were transfected with 100 ng of proviral DNA (pSVC21.BH10-*Bgl* or pSVC21.BH10-*Bgl*/MP2/MNC) and 100 ng of pBlueScript. They were simultaneously transfected with siRNAs (control, siALIX, siTSG101 or siALIX and siTSG101). After 24 hours the siRNA transfections were repeated.

2.4.6. ALIX overexpression assay

293T cells in 24-well plates were transfected with 30 ng of proviral DNA (pSVC21.BH10-*Bgl*/PTAP$^-$ or pSVC21.BH10-*Bgl*/LD4/PTAP$^-$), with or without pBJ5-ALIX-HA, made up to 200ng with pBlueScript.

2.5. Protein preparation and analysis

2.5.1. Harvesting transfected cell supernatants and lysates

Cell culture supernatants were transferred to microcentrifuge tubes and clarified by centrifuging at 14 000 rpm in a benchtop microcentrifuge for 2 minutes. The supernatant was then carefully removed without disturbing the cell pellet and stored at -80 ^{o}C.

Transfected cells were lysed with 1X Cell Culture Lysis Reagent (CCLR) (Promega) and then placed for 30 minutes on a plate rocker. Lysates were then transferred to microcentrifuge tubes.

2.5.2. Purification of virions from cell culture supernatants

Clarified supernatants were mixed with OptiPrep Density Gradient Medium (Sigma-Aldrich) in a ratio of 2:1 before centrifuging at 20 000 rpm in a TLA-55 rotor in an Optima MAX-E ultracentrifuge (Beckman Coulter). Samples were centrifuged for 90-120 minutes at 4 ^{o}C before carefully discarding the supernatant and resuspending the virus pellet in 10 µl PBS saline, and storing at -80 ^{o}C.

2.5.3. Western blotting

Samples (cell lysates or purified virions) were mixed with an equal volume of 2X Laemmli buffer and boiled for 5 minutes at 95 ^{o}C before centrifuging for 2 minutes at 14 000 rpm in a microcentrifuge. Samples were run alongside Precision Plus Dual Colour protein ladder (Bio-Rad) on 12-15% sodium

dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) gels in Western blot running buffer at 120-140 V.

To separate the 24 kilodalton (kDa) CA and 25 kDa CA-SP1 bands in provirus-transfected cell lysates 15% gels were run for 4 hours or until the 25 kDa ladder marker was about 1 cm from the edge of the gel. Proteins were transferred from the gel to a nitrocellulose membrane (GE Healthcare) by wet electrotransfer in a Mini-PROTEAN Tetra Cell (BioRad). Transfer was performed in Western blot transfer buffer at 100 V for 90 minutes on ice or 15 V at 4 ^{o}C overnight. Membranes were blocked by incubating at room temperature in 5% milk in PBS for 30 minutes, before incubating for 2 hours at room temperature or overnight at 4 ^{o}C in primary antibody (Table 7) diluted in 5% milk in PBS.

Following incubation in primary antibody, membranes were washed for 10 minutes 3 times in PBS-0.1% tween, before incubating for 2 hours at room temperature or 4 ^{o}C overnight in secondary antibody diluted in PBS (Table 7). This was followed by a repeat of the washing steps, before a 5-minute wash in PBS. Membranes were covered with 1 ml ECL (GE Healthcare) or ECL Prime (GE Healthcare) (for increased sensitivity) before exposure to X-Ray film and development using an automatic film developer.

Table 7 – antibodies used in Western blotting

Protein target	Antibody type	Species	Catalogue number, supplier	Concentration used
HIV-1 CA	Primary	Mouse	ARP313, NIBSC	1:20 000
GAPDH	Primary	Rabbit	ab9485, Abcam	1:2000
HA tag	Primary	Rabbit	14-6756, eBioscience	1:2000
ALIX	Primary	Mouse	3A9, Cell Signaling Technology	1:1000
TSG101	Primary	Mouse	ab83, Abcam	1:2000
Mouse IgG	Secondary (Horseradish peroxidase (HRP)-conjugated)	Horse	7076, Cell Signaling Technology	1:20 000
Rabbit IgG	Secondary (HRP-conjugated)	Donkey	sc-2313, Santa Cruz Biotechnology	1:20 000

2.5.4. Pulse-chase labelling

293T cells in 6-well plates were transfected with 1 µg of pSVC21.BH10-*Bgl* wild type or DIS, Stem B, Loop B or LD4 mutant proviral DNA per well. Four identical plates were prepared in total – one for each time point. 24 hours post-transfection, the medium was discarded and cells were gently washed twice with PBS, before adding 500 µl of methionine/cysteine-free DMEM (Thermo Fisher) supplemented with 4mM L-Glutamine (Thermo Fisher) and 10% (v/v) FBS. Cells were incubated at 37 ^{o}C in 5% CO_2 for 40-60 minutes to starve them of methionine and cysteine. 10 µl (≈ 4 MBq) of EasyTAG EXPRES^{35}S

protein labelling mix (Perkin Elmer) was added directly to the medium in each of the wells, and plates were incubated at 37 ^{o}C, 5% CO_2 for 40 minutes to radioisotopically label newly synthesised protein with the beta radiation emitter ^{35}S. The medium was then discarded, and the cells gently washed twice with PBS. 500 µl of DMEM supplemented with 2 mM methionine and 2 mM cysteine was added to each well. Plates for the 20, 40 and 60-minute time points were returned to the incubator, with the remaining plate providing the samples for the 0-minute time point. At each time point the supernatant was discarded and the cells were washed with 500 µl of ice-cold PBS before adding 300 µl of RIPA buffer and incubating for 5 minutes on ice. Lysates were pipetted up and down to break them up, and transferred to microcentrifuge tubes for overnight rotation at 4 ^{o}C. Following overnight rotation, the lysates were centrifuged in a microcentrifuge at 13 000 rpm, 4 ^{o}C for 15 minutes to pellet cell debris, and 500 µl of supernatant was harvested. 0.25 µl (3.25 µg) of ARP313 anti-p24 antibody (NIBSC) was added to the lysate supernatant and the samples were rotated overnight at 4 ^{o}C. After overnight rotation 50 µl of protein G-linked sepharose beads were added to each sample and rotated at room temperature for 2 hours (beads were prepared by adding 50 µl of water to 50 µl of Protein G Sepharose, Fast Flow (Sigma Aldrich), centrifuging at 6000 rpm for 2 minutes in a microcentrifuge, discarding supernatant and resuspending in 50 µl water). The samples were then centrifuged at 6000 rpm for 2 minutes in a microcentrifuge to immunoprecipitate Gag. The beads were washed by discarding the supernatant, resuspending in 500 µl of RIPA buffer, and repeating the centrifugation before again discarding the supernatant. Beads were resuspended in 50 µl of 2X Laemmli buffer and boiled for 5 minutes to separate the immunoprecipitated Gag from the beads. Centrifugation was repeated to pellet the beads and 15 µl of supernatant was loaded into a 15% SDS-PAGE gel and run at 120 V for around 4 hours to enable maximum separation of CA-SP1 and CA bands. Gels were immersed in gel fixing solution for 90 minutes to allow the proteins to form insoluble aggregates in the gel to maximise band sharpness. The solution was then discarded, and gels were immersed in 50 ml of Amplify Fluorographic Reagent (GE Healthcare) for 30 minutes on a shaker. To reduce the risk of gels cracking during the drying process the gels were immersed in 50 ml of gel drying solution for 15 minutes. Gels were then placed onto Whatman 3MM chromatography paper (GE Healthcare) and covered in cling film, before placing on a vacuum gel drier for 1 hour. Dried gels were exposed to Biomax MR film (Kodak) and the films developed using an automatic film developer.

2.5.5. Quantifying Gag processing and release

Western blots of cell lysates and virions were performed and multiple exposures of blots were made to try to obtain images with clearly defined but not over-exposed bands. Films were scanned and ImageJ software (NIH) was used to quantify the density of bands. To calculate the relative efficiency of a processing step, the ratio of the density of products to precursors in mutants was divided by the ratio in the wild type.

Similarly, to calculate the relative release ratio the sum of the densities of Gag-associated bands in mutant virion blots (often only the CA band was observed in virion blots) was divided by the sum of Gag-associated bands in their respective lysate blots and normalised against the wild type ratio.

2.5.6. Quantifying luciferase expression

At 48 hours post-transfection 293T cell supernatants were discarded and cells were lysed with 150 µl of 1X CCLR per well, before centrifuging at 14 000 rpm for 2 minutes in a benchtop microcentrifuge. 5 µl of clarified cell lysate was added per well of a medium binding 96-well flat-bottomed polystyrene plate and a GloMax luminometer (Promega) was used to inject 25 µl of Luciferase assay reagent (Promega) into each well and to record luciferase activity. To account for background, a reading from the untransfected control well was also taken.

2.6. Preparation and Northern blot analysis of viral RNA

2.6.1. Treatment to inhibit RNases

All solutions were prepared using designated RNase-free chemicals where possible and with autoclaved MilliQ water. Solutions and microcentrifuge tubes were autoclaved before use. Surfaces (pipettes, gel tanks and worktops) were sprayed and wiped with RNaseZap (Thermo Fisher).

2.6.2. Purification of RNA from virions

293T cells seeded in 6-well plates were transfected with 2.5 µg of pSVC21.BH10-*Bgl*, DIS, Stem B, Loop B or LD4 proviral DNA per well. At 48 hours post-transfection, the supernatants were harvested and processed to purify virions. 200 µl of proteinase K extraction buffer was added and samples were incubated at 37 ^{o}C for 30 minutes to digest viral proteins. Samples were then placed on ice and RNA was extracted as follows: 200 µl of 5:1 phenol:chloroform (pH 4.3-4.7) (Sigma Aldrich) was added to each sample before briefly vortexing and then centrifuging the samples at 14 000 rpm in a benchtop microcentrifuge. The upper liquid phase containing RNA was transferred to a new microcentrifuge tube and 200 µl of 24:1 chloroform:isoamyl alcohol (Sigma Aldrich) was added, and the vortexing and centrifugation were repeated, before again harvesting the upper liquid phase. RNA was precipitated by mixing in 1/10 volumes of 3 M sodium acetate pH 5.2 and then 2.5 volumes of 100% ethanol and incubating on ice for 30 minutes, followed by centrifugation at 14 000 rpm in a benchtop microcentrifuge for 20 minutes at 4 ^{o}C. The supernatant was carefully removed, and the pellet washed with 200 µl of ice cold 70% ethanol, before repeating the centrifugation step. The supernatant was discarded, and the pellet allowed to air dry for a few minutes before resuspending in 10 µl nuclease-free water. Purified viral RNA was stored at -80 ^{o}C.

2.6.3. *In vitro* transcription of Northern blot probe

A DNA probe template was first produced by PCR, using primers 1878F and 2984R-T7 (Table 2). The reverse primer contains the promoter for bacteriophage T7 polymerase, enabling *in vitro* transcription

of an ssRNA probe which is antisense to the gRNA. Nucleotides 1878-2984 are not present in spliced viral RNA, thus enabling the probe to bind to full-length gRNA specifically.

The probe was *in vitro* transcribed using the MAXIscript *In Vitro* Transcription Kit (Thermo Fisher) in the presence of digoxigenin (DIG)-conjugated UTP to enable later detection. The reaction mixture consisted of 5 µl of PCR product, 2 µl of 10X transcription buffer, 1 µl each of 10 mM ATP, CTP and GTP, 0.6 µl of 10 mM UTP, 1.14 µl of 3.5 mM DIG-UTP (Roche), and 2 µl of T7 polymerase made up to 20 µl with nuclease-free water. The reaction was incubated for 3 hours at 37 ^{o}C before twice treating with Turbo DNase by adding 1 µl to the reaction mixture and incubating for 15 minutes at 37 ^{o}C. NucAway spin columns (Thermo Fisher) were used following the manufacturer's protocol to purify the probe. The probe concentration was determined using a spectrophotometer and probes were stored at -80 ^{o}C.

2.6.4. Northern blotting of virion RNA

1 µl of 10X native gel loading solution (Thermo Fisher) was added to each 10 µl purified virion RNA sample. These were loaded alongside RiboRuler High Range RNA ladder (Thermo Fisher) in a non-denaturing 1% agarose gel. Gels were run at 4 ^{o}C in 0.5X TBE running buffer at 60 V for at least 4 hours to enable sufficient separation of dimeric and monomeric gRNA bands. Gels were visualised in a gel imaging box to check gel progress and RNA integrity (a 2:1 ratio of 28S:18S ribosomal RNA (rRNA) from the cellular RNA control indicated that RNA was not degraded in the gel). RNA was transferred from the gel onto a Hybond-N^{+} positively charged nylon membrane (GE Healthcare) by downward capillary transfer, using the apparatus setup described in the NorthernMax-Gly kit (Thermo Fisher), with 20X saline-sodium citrate (SSC) buffer. Following a 2-hour transfer at room temperature, the membrane was briefly washed in 0.5X TBE to remove excess salt and agarose, and crosslinked by baking at 80 ^{o}C for 20 minutes in a hybridisation oven. The gel was visualised under UV light in a gel imaging box to verify that transfer had been successful. The membrane was placed into a 50 ml polypropylene centrifuge tube and pre-hybridised by incubating for 30 minutes at 68 ^{o}C in 4 ml of UltraHyb Ultrasensitive Hybridisation buffer (Thermo Fisher) before addition of 1 ml of 250 ng/ml probe in UltraHyb to the pre-hybridisation solution, to give a final concentration of 50 ng/ml. Hybridisation was performed overnight at 68 ^{o}C in a roller hybridisation oven with the centrifuge tube inside the glass hybridisation tube. This meant that less solution was required, and it was easier to work with the membrane. The next day the hybridisation solution was discarded, and the membrane washed for 10 minutes at room temperature in low stringency wash buffer followed by twice washing for 15 minutes at 68 ^{o}C in high stringency wash buffer. The DIG luminescent detection kit (Roche) was used for washing and binding of the alkaline phosphatase-conjugated anti-DIG antibody, following the manufacturers' protocol. CDP-star (Thermo Fisher) was used as a substrate, before exposure of the membrane to X-ray film and development in an automatic film developer.

Dimer and monomer bands from northern blots were quantified by densitometry using ImageJ (NIH) and dimerisation efficiency was calculated by dividing the % dimer for each mutant by the % dimer for the wild type.

2.7. Assessing virion infectivity and stability

2.7.1. TZM-bl infectivity assay

293T cells seeded in a 24-well plate were transfected at 24 hours with 30 ng of pSVC21.BH10-*Bgl*, DIS, Stem B, Loop B or LD4 proviral DNA, 160 ng of pBlueScript and 10 ng of pCMV-VSV-G to produce pseudotyped viruses capable of a single round of infection. In addition, a well was left untransfected as a control.

48 hours post-transfection 10 µl of clarified supernatant (harvested supernatant centrifuged at 14 000 rpm for 2 minutes to pellet cells) was added to each well of a 24-well plate of 80% confluent TZM-bl cells to infect them. TZM-bl cells have been engineered to contain stably integrated firefly luciferase and E.coli β-galactosidase under control of the HIV-1 promoter (Wei et al., 2002). Upon infection, the HIV-1 Tat protein activates these genes enabling quantitation of infectivity. They are also engineered to express high levels of CD4 and CCR5 (Platt et al., 1998) for Env binding but in this context this is not important since the viruses have a deletion in Env and are instead pseudotyped with VSV-G, which uses the low-density lipoprotein receptor (LDLR) receptor for binding (Finkelshtein et al., 2013).

At 48 hours post-infection cell supernatants were discarded and cells were lysed with 150 µl of 1X CCLR per well, before centrifuging at 14 000 rpm for 2 minutes in a benchtop microcentrifuge and discarding the pellet. Lysates were either stored at -20 ^{o}C or were used immediately for analysis. Luciferase activity was measured as described in Section 2.5.6. Luciferase values were then divided by the input virus concentration as determined by enzyme-linked immunosorbent assay (ELISA) to give a measure of infectivity per unit virus. Finally, all values were normalised against the wild type.

2.7.2. ELISA

ELISA was used to determine the concentration of CA in cell supernatants for normalisation of luciferase readings. 96-well high binding flat-bottom polystyrene plates were coated with 25 µl per well of ELISA coating antibody solution and left overnight. The next day 1/10 volumes of 1X CCLR was added to supernatants to break down viral membranes. Supernatant samples and a recombinant p24 protein standard (Aalto Bio Reagents) were serially diluted in 0.05% Empigen in 1X Tris-buffered saline (TBS). Plates which had been coated overnight were incubated for 1 hour at room temperature with 100 µl of 5% bovine serum albumin (BSA) in 1X TBS per well to block non-specific binding, before washing 4 times with 100 µl of 1X TBS per well. The serial dilutions of the samples and standard were then loaded, and the plate was put on a plate shaker for 90 minutes, followed by a repeat of the TBS wash. 25 µl of ELISA secondary antibody solution was added to each well before shaking for a further

1 hour. Wells were washed 4 times with 100 µl of PBS-0.1% Tween per well before adding 25 µl of Lumi-Phos Plus (Lumigen) to each well and incubating at room temperature in the dark for 30 minutes. A Glomax Luminometer was used to read luminescence. A standard curve was generated and used to interpolate CA concentrations from supernatant samples.

2.7.3. Capsid stability assay

The capsid stability assay was based on a previously published protocol (Ohagen and Gabuzda, 2000). Purified virions resuspended in PBS from pSVC21.BH10-*Bgl* or LD4 mutant transfected cells were made up to 5 µl with PBS alone or with 2 µl or 4 µl of 0.05% Triton X-100 in PBS, and incubated for 10 minutes at room temperature. 45 µl of 50 mM Tris-HCl pH 7.4 was added and samples were incubated for 1 hour at 37 ^{o}C, before centrifuging for 1 hour at 14 000 rpm in a microcentrifuge to pellet intact particles. The supernatant was carefully harvested leaving 5 µl which was mixed to resuspend the pellet. Samples were run on an SDS-PAGE gel before Western blot detection with ARP313 anti-p24 antibody and quantitation of the density of the pellet and supernatant bands. The proportion of p24 pelleted out of total p24 for each Triton X-100 treatment was calculated.

2.7.4. Statistics

A two-tailed two-sample unpaired Student's t-test was used for statistical analysis of population means. This test carries several assumptions including:

- The scale of measurement is continuous
- The data follow a normal distribution
- The variances of each population are approximately equal

The t-test is ideal for the type of data generated in this project (e.g. comparing mutant and wild type Gag processing) as it is limited to two categories of data – it is mathematically identical to the one-way ANOVA with two categories – and the assumptions can be reasonably made. Even if data do not follow a normal distribution, the t-test is not sensitive to this providing the distributions of the two groups are similar (McDonald, n.d.).

In some cases a one-sample test was used when the mean value obtained from a random sample was to be compared against a theoretical value.

In both types of test, the difference between means/between the mean and expected value was deemed to be statistically significant if the p value was below 0.05.

When plotting graphs with $n \geq 3$ for all values, where n represents the number of independent experiments performed, the error bars plotted represent standard deviation. In graphs where $n < 3$ for some or all values, the error bars represent range.

In all graphs the values plotted represent the mean of n independent experiments.

3. Phenotypic effects of SL1 mutations

3.1. Introduction

3.1.1. Importance of SL1 for virus replication

Given its importance in facilitating gRNA dimerisation and packaging, it is not surprising that SL1 mutants are often severely attenuated (selected examples in Table 8 with further examples in Table 11 in Appendix D). The majority of nucleotides in the stems and internal loops of SL1 are 100% conserved across all sequenced viral isolates (Greatorex et al., 2002), and analogues of SL1, featuring palindromic loops, are also found in HIV-2 and SIV (Berkhout, 1996).

Despite being severely replication-defective, many SL1 mutants have only moderate defects in dimerisation and packaging (Table 8). This suggests SL1 mutation affects other steps of the virus replication cycle, either unrelated to or because of defective dimerisation and packaging. For example, the later steps of reverse transcription have been shown to be affected by large SL1 mutations (Paillart et al., 1996; Shen et al., 2000), likely because of a destabilising effect on the structure of the Pbs.

Table 8 – phenotypes of selected SL1 mutants

Mutant name	Infectivity (% WT) (Cell type)	Dimerisation (% WT) (Cell type)	Packaging (% WT) (Cell type)	Reference
Large deletions				
Δ248-261	< 1% (MT-4)	57% (COS7)	57% (COS7)	(Laughrea et al., 1997)
LD3/Δ241-256	< 1% (MT-2)	53% (COS7)	20% (COS7) 30% (COS7)	(Liang et al., 1998; Shen et al., 2000)
LD4[1]	< 1% (MT-2)	-	70% (COS7)	(Liang et al., 1999b, 1998)
ΔSL1	< 5% (PM-1)	56% (293T)	50% (293T)	(Ristic and Chin, 2010)
Palindrome mutants				
G-loop	< 10% (HOS)	-	83% (293T)	(Clever and Parslow, 1997)
S257-259	1% (SupT1)	-	100% (COS7) 50% (SupT1)	(Paillart et al., 1996)
GGCG[2]	< 5% (MT-4)	59% (COS7)	60% (COS7) 78% (COS7)	(Laughrea et al., 1999; Shen et al., 2000)
Stem mutants				
Δ243-247	< 0.1% (MT-4)	56% (COS7)	60% (COS7)	(Laughrea et al., 1999; Shen et al., 2000)
Δ248-256	< 0.1% (MT-4)	53% (COS7)	70% (COS7)	(Laughrea et al., 1999; Shen et al., 2000)
Δ243-246[3]	< 0.1% (COS-7)	59% (COS7)	25% (COS7)	(Shen et al., 2001)
Δ274-277	< 0.1% (COS-7)	59% (COS7)	40% (COS7)	(Shen et al., 2001)
Loop mutants				
Δ271-273[4]	< 1% (COS-7)	59% (COS7)	44% (COS7)	(Shen et al., 2001)
Δ247	< 5% (COS-7)	59% (COS7)	65% (COS7)	(Shen et al., 2001)

Numbers in superscript format indicate mutants that are equivalent to the LD4 (1), DIS (2), Stem B (3) and Loop B (4) mutants used in this study. Mutant names that include nucleotide positions use the numbering system shown in ***Figure 3*** *referring to position in the transcribed RNA, and not the system used in Materials and Methods which refers to position in the provirus.*

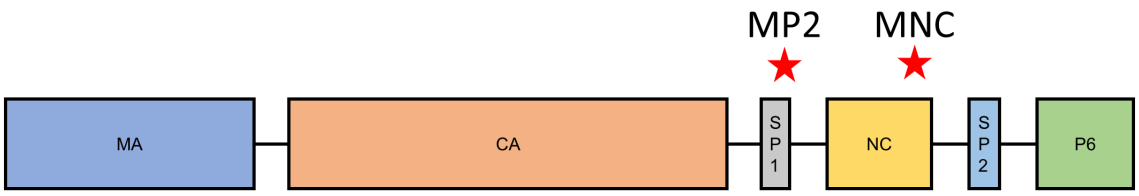
3.1.2. Compensatory mutations in SL1 mutants

Virus evolution studies can provide valuable information on the interactions taking place between different regions of the viral genome, to identify novel functions associated with those regions. Sequencing of infectious viruses that emerge from cultures of attenuated mutants can reveal

compensatory mutations. These are mutations that partially or fully restore a phenotype disrupted by a mutation without correcting the underlying mutation.

After 18 passages of the severely defective LD3 SL1 mutant, Liang *et al* identified a revertant with near wild type infectivity and packaging. The initial deletion in SL1 remained, but two compensatory mutations in Gag termed MP2 and MNC appeared at the 12th residue of the SP1 domain (also known as p2) and the 24th residue of the NC domain respectively (Liang et al., 1998) (Figure 13). Both mutations involve a threonine to isoleucine substitution. The same compensatory mutations also appeared during culture of mutants with deletions between the Pbs and SL1 (Liang et al., 2000) and their introduction into the LD4 SL1 mutant also restored replication to near wild type levels (Liang et al., 1998).

Figure 13 – Location of compensatory mutations in Gag



Cartoon diagram of Gag showing position of the MP2 and MNC compensatory mutations in the context of its domains (drawn to scale). The MP2 mutation is a T→I mutation at position 12 of the SP1 domain, and the MNC mutation is a T →I mutation at position 24 of the NC domain.

The nearly fully infectious LD3-MP2-MNC mutant was passaged a further 11 times and two additional mutations which supplement the original compensatory mutations were identified in MA and CA and termed MA1 and CA1 (Liang et al., 1999b).

Despite altering both gRNA and Gag amino acid sequence, the compensatory mutations act at the protein level because LD3-MP2-MNC gRNA could not compete for encapsidation into wild type virions when provided *in trans*. Substitution of T12 in SP1 and T24 in NC for every other amino revealed only that valine, leucine, and methionine, which have long hydrophobic sidechains, could functionally replace the isoleucine residue to restore infectivity to 5'UTR mutants (Rong et al., 2001). Wild type virus infectivity was greatly reduced when T12 in SP1 was replaced with either aspartic acid or glutamic acid, which have negatively charged side chains, or tryptophan which has a bulky hydrophobic side chain. This suggests that the mechanism of rescue is to restore a protein-protein or protein-RNA interaction.

The residue altered by the MNC mutation is located immediately downstream of the histidine residue which is a critical component of the first zinc finger of NC. This mutation alone substantially restores packaging to 5'UTR mutants (Liang et al., 2000, 1998; Shen et al., 2000) but has no effect on wild type

virus packaging (Rong et al., 2001). The BH-D2 mutant, which has nucleotides in the Pbs hairpin deleted, can be rescued by either MNC or by a G to A substitution in U5, which would strengthen the U5:AUG interaction to promote formation of the packaging-competent conformer (Figure 6 right). These data strongly suggest that the role of MNC is to restore a defective interaction between NC and gRNA.

The function of MP2 is less clear. As it is located in SP1, which must be cleaved from CA for virion maturation to occur (Wiegers et al., 1998), the Gag processing properties of large SL1 mutants were examined. Intriguingly these mutants exhibited a delay in the proteolytic cleavage of SP1 from its precursor CA-SP1 (Liang et al., 1999a), resulting in the production of virions with an immature morphology. Both compensatory mutations were required to correct this defect (Shen et al., 2000) suggesting that delayed CA-SP1 processing contributes to the severe attenuation of these mutants.

Like NC, SP1 is also involved in packaging (Kaye and Lever, 1998). The MP2 mutation has been shown to restore packaging specificity – the ability to discriminate between gRNA and spliced vRNA – to SL1 (Russell et al., 2003) and NC (Roy et al., 2006) mutants, although unlike MNC it does not increase packaging efficiency.

The location of MP2 in SP1 suggests that its purpose is to correct the CA-SP1 processing defect that results from SL1 mutation.

3.1.3. Possible mechanisms for the link between SL1 mutation and delayed CA-SP1 processing

The finding that SL1 mutation delays CA-SP1 processing raises the possibility that efficient genome dimerisation is required for optimal processing of Gag and production of mature, infectious viral particles. One potential mechanism is that GagPol molecules utilise gRNA dimers as a scaffold upon which to dimerise, to allow the self-excision of PR from its precursor GagPol. A prediction arising from this hypothesis would be that all steps of Gag processing are delayed. Alternatively, the effects of SL1 mutation on dimerisation and CA-SP1 processing may be unrelated. MP2 and MNC cannot correct defective gRNA dimerisation in the LD3 mutant (Shen et al., 2000) showing that efficient Gag processing is possible in the absence of efficient dimerisation. Another possibility is that the effect on processing is due to the loss of important Gag binding sites (Abd El-Wahab et al., 2014; Damgaard et al., 1998; Kenyon et al., 2015; Wilkinson et al., 2008) affecting a crucial interaction.

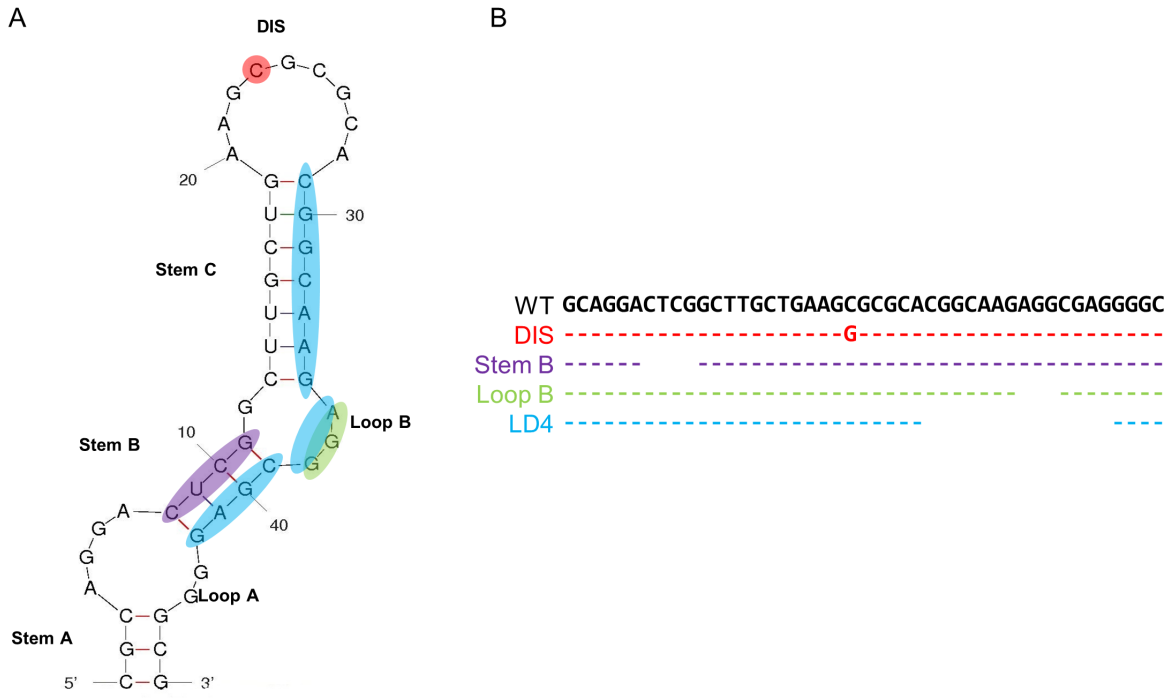
In HIV-2, mutation of a GGAG motif critical for gRNA dimerisation, at the base of SL1, resulted in the formation of virions with abnormal morphology (L'Hernault et al., 2007) caused by delayed processing of MA-CA-p2 (L'Hernault et al., 2012). Like in HIV-1, a point mutation (MA 70TI) arose in long-term culture, however unlike in HIV-1 this mutation rescued gRNA dimerisation in addition to Gag processing (L'Hernault et al., 2012).

3.2. Experimental aims

Studies on the effect of SL1 mutation on Gag processing have only been performed using mutants containing large deletions in SL1. The LD3, LD4 and LD5 mutants involve 16, 14 and 30 nucleotide deletions respectively (Liang et al., 1999a). Such large deletions are likely to have a significant impact on the structure and function of neighbouring RNA motifs (Figure 3) and may also affect the balance between the two conformers of the 5'UTR (Figure 6). This confounds interpretation of whether the dimerisation and processing defects are related or are independent consequences of widespread disruption of the 5' UTR.

Therefore, for this study, 3 smaller mutants were studied to limit disruption to SL1 and minimise the impact on neighbouring structures. The DIS mutant contains a single nucleotide substitution in the DIS palindrome, the Stem B mutant a 4-nucleotide deletion in the stem of SL1 and the Loop B mutant a 3-nucleotide deletion in the internal loop (Figure 14).

Figure 14 – Structure of SL1 and mutations introduced



A) Secondary structure prediction of SL1, with stems and loops labelled. The nucleotides deleted or modified in the mutant proviruses are shaded with colours corresponding to the mutant provirus. B) Proviral DNA sequence of SL1, from HXB2 reference sequence, accession number K03455.1, and modifications made in the mutant viruses. Dashes represent nucleotides that are identical to the wild type, blanks represent nucleotides that have been deleted and letters represent nucleotides that have been substituted. The colour coding used corresponds with the shading in (A).

To establish if a correlation exists between gRNA dimerisation and Gag processing, these phenotypes were assessed in each mutant using Northern blotting, pulse-chase and Western blotting experiments. The relative infectivity of each mutant was also assessed, and compensatory mutations were introduced into the wild type and mutant proviruses to assess their effect on infectivity and Gag processing.

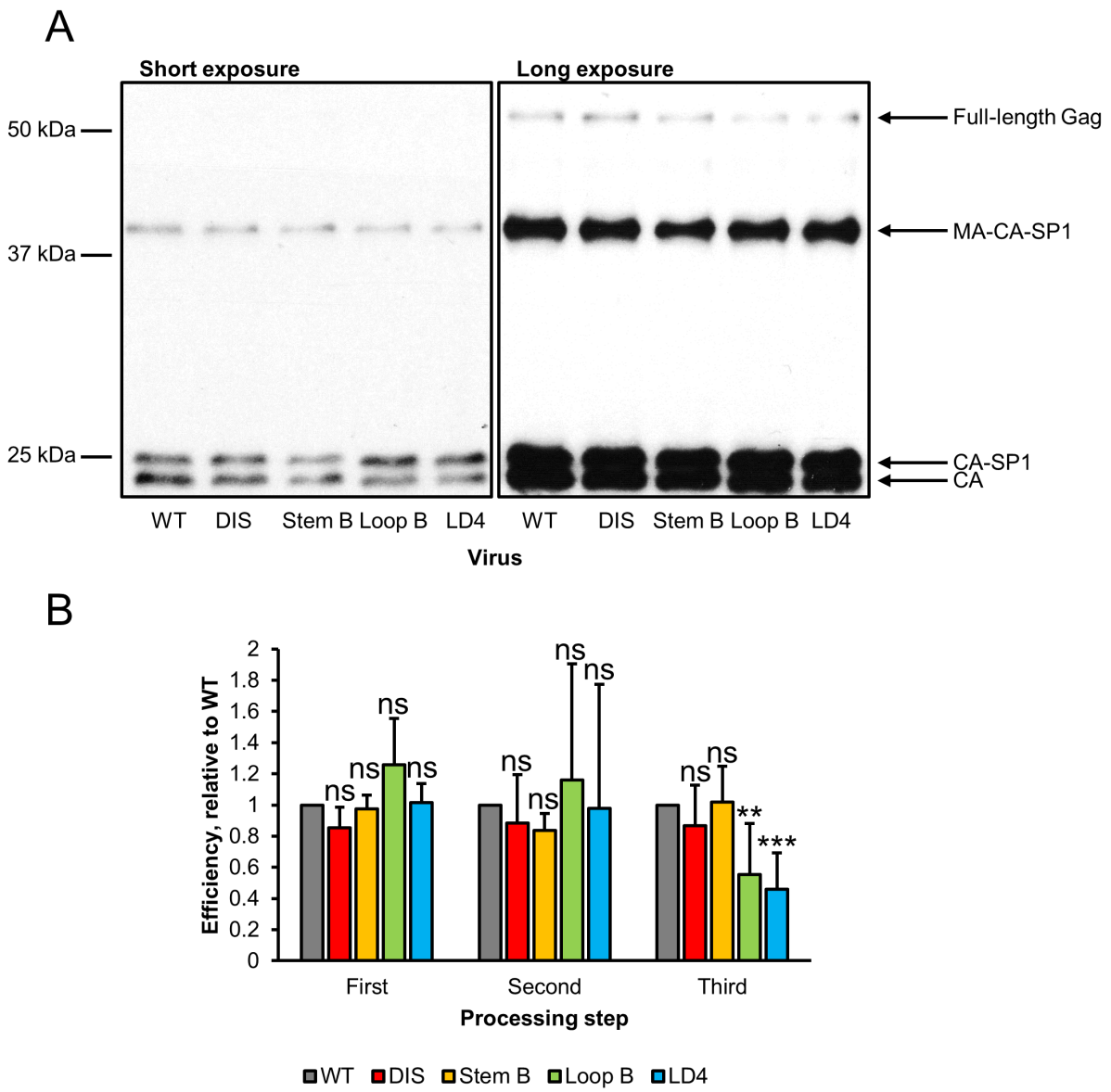
3.3. Results

3.3.1. Analysis of Gag processing

Preliminary analysis of the effect of SL1 mutations on Gag processing was performed using pulse-chase metabolic labelling, to observe Gag processing over time (see Appendix C for experimental setup and results). Due to high levels of background and low contrast of some bands the experiment provided qualitative to semi-quantitative results at best and can only be regarded as supporting evidence, however visual inspection of the blots seemed to indicate that only the loop B and LD4 mutants possess a CA-SP1 processing defect (for further discussion see Appendix C).

To corroborate the results of the pulse-chase experiments, cells transfected with provirus constructs were harvested 48 hours post-transfection and lysates were analysed by Western blotting (Figure 15). Unlike the pulse-chase experiment which shows the kinetics of processing over time, Western blotting provides a snapshot of the state of Gag processing in the cell at a particular time point. The lack of background from cellular proteins on Western blots enabled quantitation of the efficiency of each of the Gag processing steps (Figure 12) – Full-length Gag → MA-CA-SP1 (first cleavage), MA-CA-SP1 → CA-SP1 (second cleavage) and CA-SP1 → CA (third cleavage) – by comparing the ratios of processing products to their precursors.

Figure 15 – Gag processing intermediates in SL1 mutants observed by Western blotting



*A) Western blot of cell lysates following transfection with the indicated proviruses. Short and long exposures of the blot are shown to show each Gag cleavage product. The identities of each of the Gag processing intermediates is indicated on the right. B) The graph shows the mean efficiency of each processing step relative to wild type from at least 3 independent experiments. The efficiency of each step was determined by calculating the ratio of products to substrates (for example, for the first step the ratio of MA-CA-SP1 to full-length Gag was calculated). Error bars represent standard deviation. In this and all other figures (unless otherwise stated) a two-tailed Student's t-test was performed (ns = 'not significant', * = 'p<0.05', ** = 'p<0.01', *** = 'p<0.001').*

In agreement with the pulse-chase data, both the Loop B and LD4 mutants displayed an accumulation of CA-SP1 relative to the wild type, whilst the DIS and Stem B mutants exhibited wild type CA-SP1 processing (Figure 15). There were no defects in the preceding steps of processing for any of the mutants. However, the use of a weak detergent to lyse cells in preparation for Western blotting

(1XCCLR was used, which contains 1% Triton X-100) may have resulted in incomplete breakdown of immature virus particles at the cell membrane. This may explain the weak full-length Gag band in Figure 15. Whilst this could affect quantitation of the absolute rate of the first step of processing, it is unlikely to affect the interpretation of the results presented here as the relative rates of processing between the mutants and wild type are of interest in this context.

These results show that the Gag processing defect is only caused by particular SL1 mutations and is specific to the final stage of cleavage, a step critical for maturation (Figure 12).

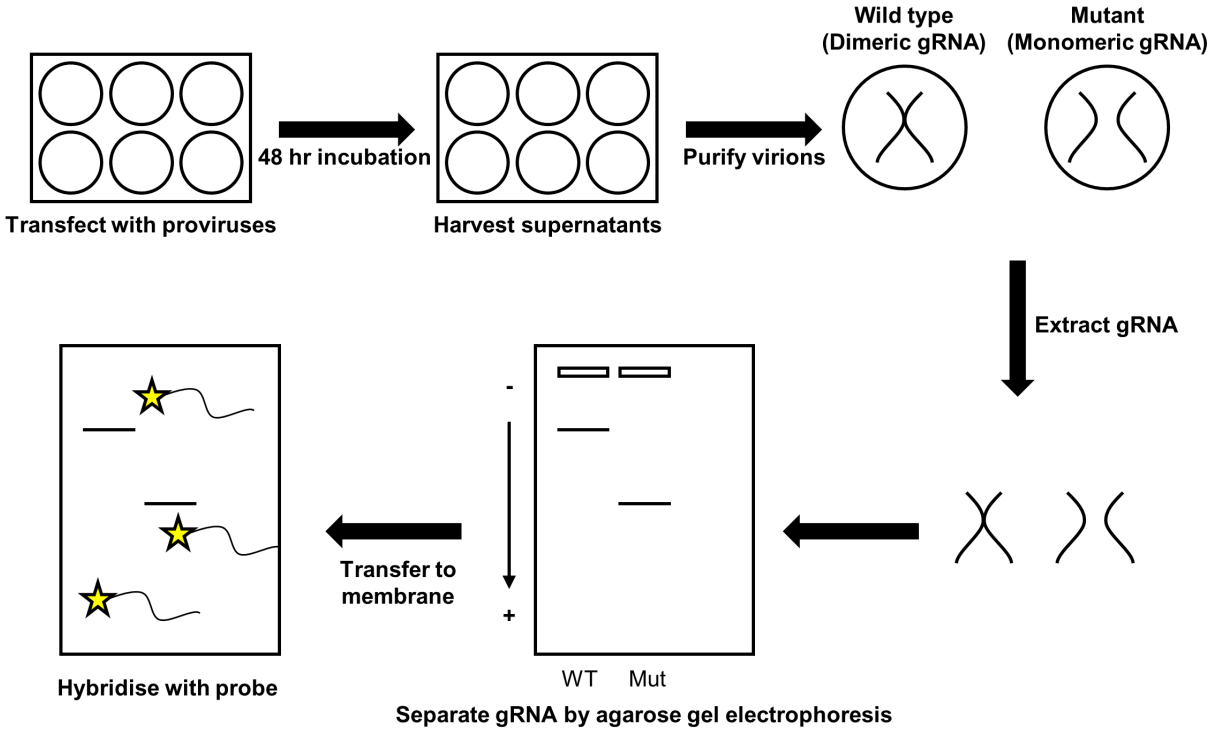
3.3.2. Effect of SL1 mutations on gRNA dimerisation

To determine whether the efficiency of CA-SP1 processing correlates with gRNA dimerisation, Northern blotting was performed on gRNA isolated from virions (Figure 16). gRNA bands from wild type and mutant virions were quantified by densitometry to enable measurement of dimerisation and packaging efficiencies. Packaging efficiency was calculated by dividing the intensity of the total gRNA bands by the concentration of input virus, determined by ELISA to give an approximate measure of the amount of gRNA packaged per unit of CA (Figure 17).

Despite disrupting the palindrome at the tip of the DIS by changing its sequence from GCGCGC → GGGCGC, gRNA from DIS mutant virions consistently ran with the same profile as the wild type (Figure 17A lane 2), suggesting that a perfect palindrome is not required for efficient genome dimerisation. On the other hand, gRNA extracted from the Stem B and Loop B mutants (Figure 17A lanes 3 & 4) was much more monomeric, suggesting that there is reduced formation of dimers, or formation of dimers with reduced stability that dissociate under the conditions used in this experiment. The finding that only the Loop B mutant exhibits a CA-SP1 processing defect, despite the Stem B mutant also being deficient in dimerisation, suggests that dimerisation is not required for efficient CA-SP1 processing.

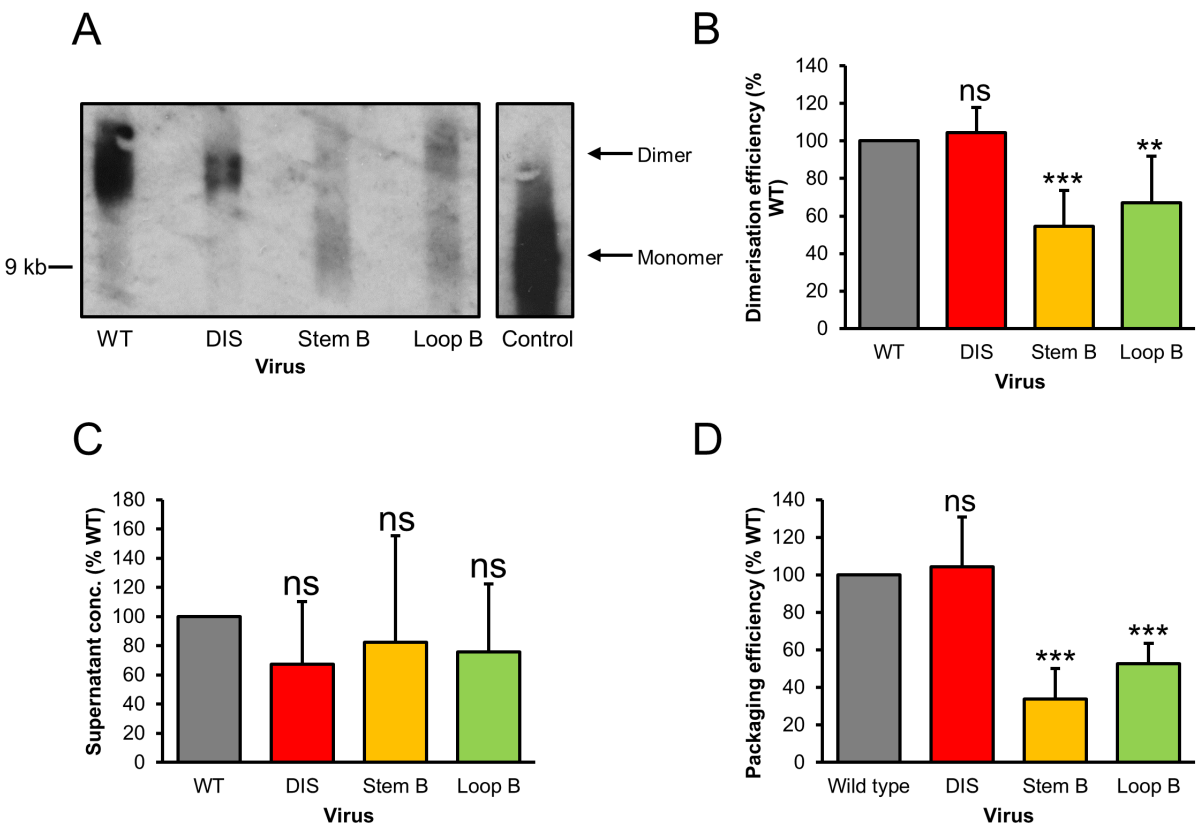
The weakness of the intensity of the Stem B and Loop B bands is consistent with these mutants having more severe packaging defects (Figure 17**D**) as previously shown (Shen et al., 2001). There was a good correlation between the ability of gRNA to form dimers and the efficiency of packaging (Figure 17) which is consistent with the notion that genome dimerisation is required for efficient packaging to occur. The DIS mutant had wild type like dimerisation and packaging, whereas the Stem B and Loop B mutants were defective in both phenotypes. Packaging of DIS mutant RNA in the blot presented appears lower than wild type but this was not observed in the other blots performed.

Figure 16 – Northern blot experimental setup



Supernatants from 293T cells transfected with proviral DNA were clarified and virions were purified, followed by digestion of the viral proteins and purification of RNA. Dimers and monomers were separated by agarose gel electrophoresis, and then transferred to a membrane for hybridisation with a DIG-labelled antisense probe specific to full length gRNA. All procedures were carried out at 4 ^{o}C to minimise degradation of loose dimers. The presence of DIG in the probe enabled capture of an alkaline phosphatase-conjugated anti-DIG antibody, followed by addition of substrate and exposure of the membrane to film.

Figure 17 – Genome dimerisation and packaging in SL1 mutants



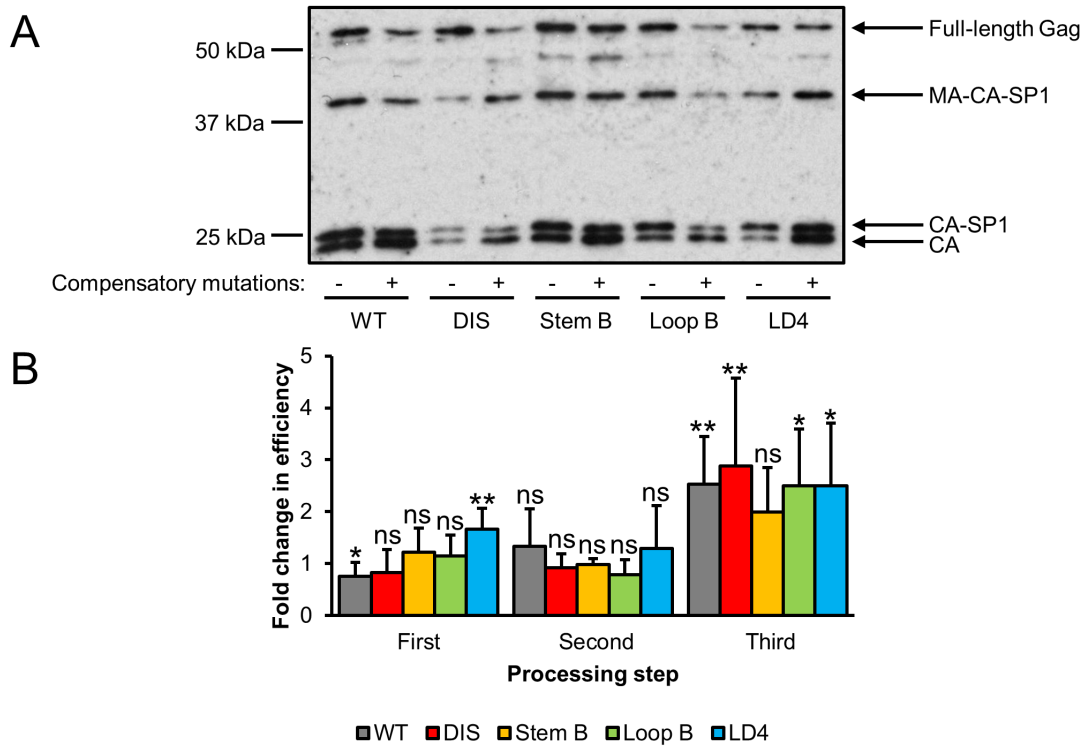
A) Northern blot of purified viral RNA hybridised with a gRNA-specific probe. The positions of the dimer and monomer bands are indicated. Viral RNA isolated from the cytoplasm was run in a control lane to verify that the lower band represents the monomeric species. B) Quantitation of dimerisation efficiency relative to wild type, calculated as % dimer in the mutant divided by % dimer in the wild type. The graph shows the mean values from at least 6 independent experiments were performed. Error bars represent standard deviation. C) ELISAs were performed on supernatants to determine p24 concentration. At least 4 independent experiments were performed. D) Quantitation of packaging efficiency relative to wild type. The amount of gRNA in each virion prep was quantified by performing densitometry on Northern blot bands, before dividing by the average concentration of virus in supernatants as determined by ELISA. The graph shows the mean values from 4 independent experiments. Error bars represent standard deviation.

3.3.3. Indiscriminate acceleration of CA-SP1 processing by compensatory mutations

Given the finding that deletion of 3 nucleotides from the Loop B motif of SL1 replicates the CA-SP1 processing defect observed in the LD4 mutant, it was of interest to see if the compensatory mutations could also restore processing in the Loop B mutant.

The MP2 and MNC mutations were introduced into the wild type and mutant provirus constructs, before transfection into 293T cells and Western blotting to quantify processing efficiency (Figure 18).

Figure 18 – Effect of MP2/MNC compensatory mutations on the stages of Gag processing



*A) Western blot of cell lysates following transfection with the indicated proviruses. B) Quantitation of the fold-change in the efficiency of Gag processing (as calculated in **Figure 15**) upon introduction of compensatory mutations. A one-sample t test was performed to compare the fold-changes against the null hypothesis (fold-change of 1). The graph shows the mean values from at least 3 independent experiments. Error bars represent standard deviation.*

The compensatory mutations rescued defective CA-SP1 processing in the Loop B mutant in addition to the LD4 mutant. Interestingly, CA-SP1 processing was accelerated 2 to 3-fold in all viruses, including the wild type (Figure 18). There were two small but statistically significant changes in the first step of processing – a decrease in the wild type virus and an increase in the LD4 mutant. Lanes 3 and 4 of the gel shown suggest an acceleration of the first step of Gag processing in the DIS mutant, however this was not observed in any of the other repeats of this experiment. As noted in Section 3.3.1, the full-length Gag bands are likely to be less intense than they could be because of the weak detergent used to lyse cells, but this should not affect comparison of the relative rates of processing between mutants.

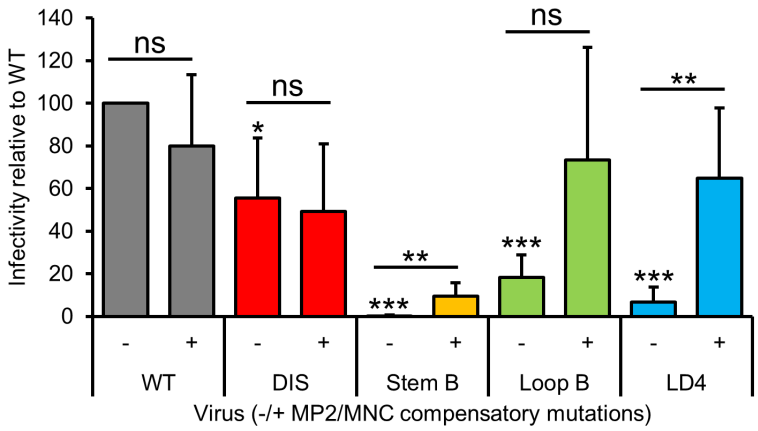
The introduction of compensatory mutations had no effect on the second step of processing, and minor effects on the first step, as discussed above, so they therefore appear to specifically accelerate the final step of processing. This occurs in an indiscriminate manner regardless of the SL1 mutation present, i.e. their effect is the same whether SL1 is fully intact or contains a small or large deletion.

3.3.4. Infectivity of SL1 mutants and effect of compensatory mutations

Virus infectivity is particularly sensitive to alterations in the timing of Gag processing, so the effect of SL1 mutations and compensatory mutations on virus infectivity was assessed.

Infectivity was determined using TZM-bl cells, which express luciferase under the control of a tat-inducible promoter. They were infected with VSV-G pseudotyped viruses, and luciferase expression was measured. To account for variations in virus input, the luciferase reading was divided by the p24 concentration, as determined by ELISA, to give a measurement of per-virion infectivity. The results are shown in Figure 19.

Figure 19 – Effects of SL1 mutations and compensatory mutations on virus infectivity



The graph shows the mean values for SL1 mutant infectivity, without and with compensatory mutations relative to wild type, from 4 independent experiments. Error bars represent standard deviation. Inter-mutant (i.e. between SL1 mutant and wild type) statistical significance is indicated on top of bars, whilst intra-mutant (i.e. between mutant with and without compensatory mutations) is indicated on top of lines between bars.

There were large differences in the infectivity of SL1 mutants (Figure 19). The DIS mutation caused a less than 2-fold reduction in infectivity, whereas the Stem B mutation produced a virus which consistently gave luciferase readings identical to negative controls, indicating a completely non-infectious virus. The Loop B and LD4 mutants had an intermediate phenotype, as some infectivity was observed albeit with an efficiency of 5 to 10-fold less than wild type.

The magnitude of the reduction in infectivity of the Stem B and Loop B mutants is greater than the magnitude of their respective packaging and dimerisation defects, indicating that other steps are affected. Likewise, the reduction in infectivity of the DIS mutant cannot be explained in terms of dimerisation and packaging as these were not affected by the mutation (Figure 17).

In agreement with a previous report (Liang et al., 1998), the compensatory mutations resulted in a statistically significant increase of LD4 mutant infectivity to near wild type levels. The mean infectivity

of the Loop B mutant was also increased to similar levels, although greater variability meant that the difference was not statistically significant. The ability of the compensatory mutations to largely restore Loop B and LD4 mutant infectivity suggests that delayed CA-SP1 processing is the major factor responsible for attenuation of those mutants.

There was also a small but statistically significant increase in Stem B mutant infectivity, turning a non-infectious virus into one capable of replication, albeit with a 10-fold reduction in infectiousness relative to wild type.

Surprisingly, despite accelerating CA-SP1 processing above its natural rate, the infectivity of the wild type virus was not affected, nor was the infectivity of the DIS mutant.

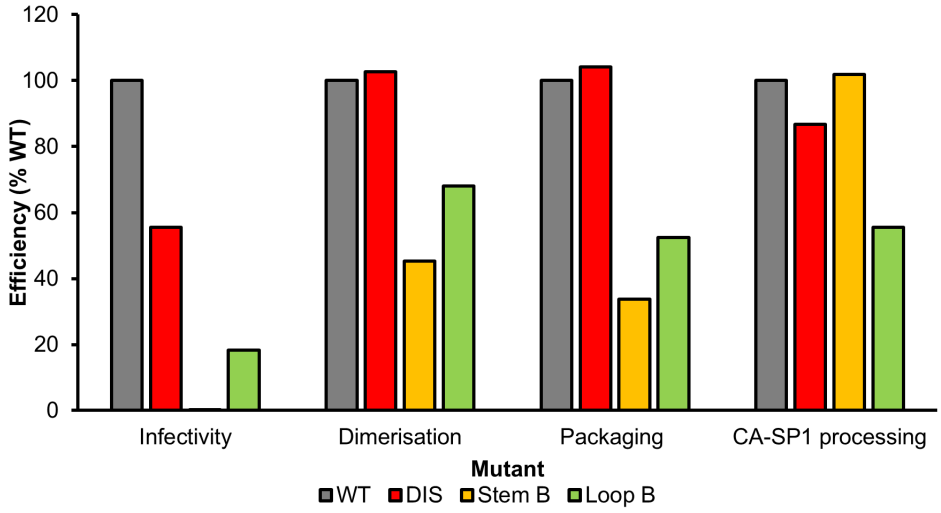
3.4. Discussion

To gain insight into the elements of SL1 required for efficient Gag processing, three SL1 mutants (DIS, Stem B and Loop B) that had not previously been assessed for their Gag processing capability were studied alongside the LD4 mutant, which served as a positive control. Gag processing efficiency, genome dimerisation, packaging, infectivity, and the effect of compensatory mutations was assessed.

3.4.1. Lack of correlation between gRNA dimerisation and Gag processing

The results from the pulse-chase and Western blotting analysis show that the Gag processing defect caused by the 14 nucleotide LD4 deletion was recapitulated only in the Loop B mutant (Figure 15). Notably, CA-SP1 processing was unaffected in the Stem B dimerisation mutant (Figure 17), demonstrating that dimerisation is not required for efficient Gag processing (summarised in Figure 20). This supports the finding by Shen *et al* that compensatory mutations could restore Gag processing to the LD3 mutant without rescuing genome dimerisation (Shen et al., 2000).

Figure 20 – Comparison of SL1 mutant phenotypes relative to wild type



Values shown are mean values compiled from previous figures in this chapter for comparison purposes.

The effect on Gag processing was specific to the final step, CA-SP1 cleavage (Figure 15), which is inconsistent with the hypothesis that genome dimerisation facilitates Gag processing by providing a scaffold for GagPol molecules to dimerise on. If this were the case, all steps of Gag processing would be expected to be delayed by mutations which disrupt dimerisation.

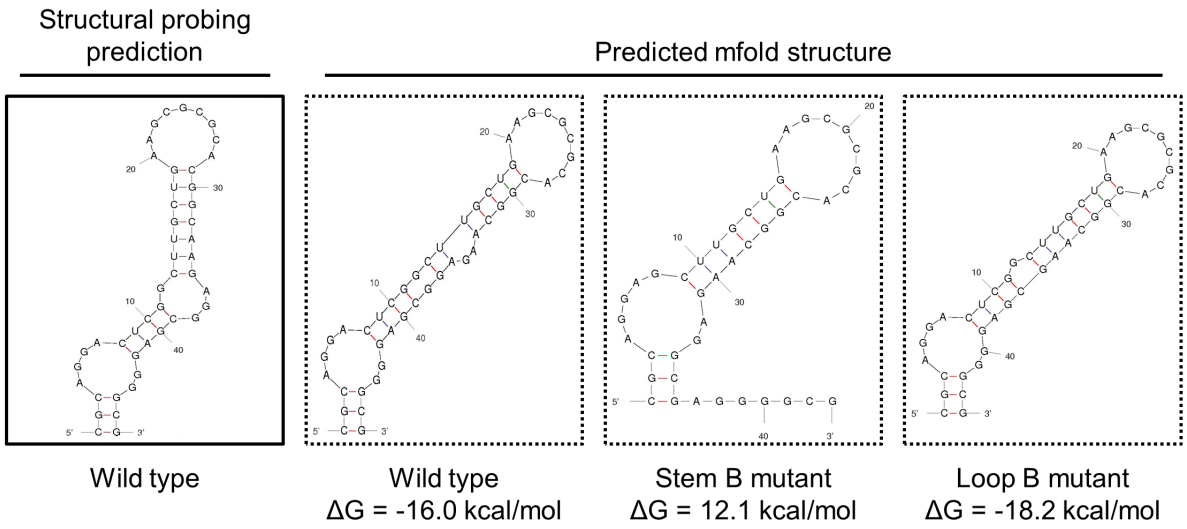
3.4.2. Possible mechanisms for Loop B involvement in Gag processing

In addition to its effect on Gag processing, mutation of Loop B is as detrimental to virus infectivity as the LD4 mutations (Figure 19). It consists of four residues – a single unpaired G on the 5' side and an unpaired AGG triplet on the 3' side – which disrupt the otherwise perfect base pairing in the main stem of SL1 (Figure 21 left panel).

The loop is thermally and conformationally unstable – a single structure could not be resolved on NMR – and its complete conservation suggests that this flexibility is important for its function (Greatorex et al., 2002). The presence of an internal hinge may serve to facilitate unwinding of SL1 to enable maturation of the kissing loop dimer to an extended dimer (Figure 7). This hypothesis is supported by *in vitro* data; NMR analysis of a 39mer RNA corresponding to SL1 showed that Loop B had a loosening effect on neighbouring base pairs in the stem (Takahashi et al., 2000b). Maturation of the 39mer into a tight dimer in the presence of NC or following incubation at 55 ^{o}C was blocked when Loop B was deleted (Takahashi et al., 2000a).

RNA secondary structure calculation software predicts that the Loop B mutation (deletion of the AGG triplet) would result in merging of Stem B and Loop B to produce a much more thermally stable 11-base pair stem containing a single unpaired base in place of Loop B (Figure 21).

Figure 21 – Predicted secondary structures of Stem B and Loop B mutants



Predicted structures and free energy changes of wild type and mutant SL1 sequences, obtained using mfold, with no constraints on base pairing. The predicted wild type structure differs slightly from the structure determined by structural probing techniques, so the predictions for the mutants serve as a best estimate.

The Stem B mutation on the other hand is predicted to create a larger internal loop (Figure 21). This would have the opposite effect to Loop B mutation on the thermal stability of SL1. The DIS mutation involves substitution of an unpaired residue so is unlikely to affect stability.

One possibility therefore is that CA-SP1 processing is linked to maturation of the dimer, with Loop B playing a critical role by weakening the stem to allow unwinding of the kissing-loop dimer. It has been speculated that the large structural rearrangement of the RNA which takes place during dimer maturation may help to coordinate virion maturation (Takahashi et al., 2000b). The changing footprint of Gag/NC on the RNA during dimer maturation may trigger a structural change in the CA-SP1 cleavage junction, facilitating processing.

A second related possibility is that Loop B mutation disrupts Gag binding to gRNA. RNA footprinting (Damgaard et al., 1998), SHAPE (Wilkinson et al., 2008) and crosslinking-SHAPE (XL-SHAPE) (Kenyon et al., 2015) all highlight Loop B and surrounding residues as a Gag binding site. More recently, Abd El-Wahab *et al* argued that Loop B is the key motif responsible for specific packaging of gRNA into virions. Deletion of Loop B significantly reduced binding of full-length recombinant Gag to an *in vitro* synthesised RNA corresponding to the first 600 nucleotides of the HIV-1 genome (Abd El-Wahab et al., 2014).

The two features which set the Loop B mutation apart from the Stem B and DIS mutants are that it appears to over-stabilise SL1 and removes a potential Gag binding site. At this stage it is not possible

to conclude if either or both of these functions is important for regulating Gag processing, but as both processes involve binding of Gag it seems plausible that binding could trigger the structural change in the CA-SP1 junction to enable cleavage by PR.

3.4.3. Correlations between dimerisation, packaging and infectivity

There was a clear correlation between dimerisation and packaging (Figure 17) consistent with the model of preferential packaging of dimers in virions (Russell et al., 2004).

These phenotypes correlated with virus infectivity (summarised in Figure 20), although the reductions in packaging alone (3-fold for Stem B and 2-fold for Loop B) were not as severe as the reductions in infectivity (50-fold for Stem B and 6-fold for Loop B), further highlighting that steps other than packaging and dimerisation may be affected by mutation of SL1. The complete loss of infectivity from the Stem B mutant may be explained by the apparent 15-fold reduction in reverse transcription efficiency (Shen et al., 2000) together with the 3-fold reduction in packaging observed in this study.

The above arguments assume that packaging, dimerisation and reverse transcription have additive effects on infectivity, but this may not be the case. Disruption to multiple stages of virus replication may result in synergism. In addition, the immediate effect on infectivity observed in this single-cycle replication assay may be different to the cumulative effect that would be observed in a multi-round assay, or in an infected individual.

Despite these uncertainties, the CA-SP1 processing defect in the Loop B mutant is likely to be a key additional contributor to reduced infectivity, because delayed processing of CA-SP1 hinders the formation of mature infectious virions (Wiegers et al., 1998).

Despite the preference for packaging gRNA dimers, monomeric gRNA bands did appear on Northern blots, especially in the Stem B and Loop B mutants, suggesting that monomeric RNA can be packaged. However, these may be the result of dimers dissociating. Despite the ambiguity associated with using Northern blotting to study dimerisation, the technique is still widely used and the results interpreted as a proxy for the efficiency of dimerisation in virions - all the dimerisation efficiency values in Table 8 and Table 11 are derived from quantitation of Northern blots. One way to exclude the presence of tight dimers is to introduce an inactivating mutation in PR to prevent dimer maturation (Rong et al., 2003) but this still does not distinguish between packaged monomers and monomers arising from loose dimers dissociating. Techniques to dual label gRNA and visualise individual molecules in cells (Chen et al., 2009; Ferrer et al., 2016) may help to answer questions about the effect of mutations on dimerisation without the caveats associated with Northern blotting.

3.4.4. The effect of the DIS mutation on dimerisation and packaging – cell type dependence?

Surprisingly, mutation of the DIS palindrome at the tip of SL1 had no effect on gRNA dimerisation or packaging (Figure 17) despite a 2-fold reduction in infectivity (Figure 19). This may be due to the palindrome playing an unappreciated role in the early stages of virus replication, which would be consistent with the observation that some artificial palindrome sequences are highly detrimental to infectivity and are never observed in nature (Laughrea et al., 1999).

There is disagreement in the literature on the phenotypic effects of DIS palindrome mutations (Table 9). Mutation resulted in reduced dimerisation (Shen et al., 2000) and packaging (Laughrea et al., 1999; Shen et al., 2000) in COS7 cells, and deletion almost completely abolished infectivity in MT-4 cells (Laughrea et al., 1999), however, as shown here, in 293T cells dimerisation was indistinguishable from wild type (Figure 17) (Hill et al., 2003; Song et al., 2007).

Phenotypic discrepancies between cell types are also observed following more extensive SL1 mutation. Despite causing severe replication defects in SupT1 cells, almost-complete deletion of SL1 had no effect on replication in peripheral blood mononuclear cells (PBMCs) (Hill et al., 2003). The researchers created two SL1 mutants, by replacing the 39-nucleotide SL1 with its natural GCGCGC palindrome, or with an artificial ACGCGT palindrome. In SupT1 cells replication was negligible (1,000 to 10,000 times less infectious than WT), however in PBMCs similar replication kinetics were observed to WT, albeit with a slight delay. This was not due to the emergence of compensatory mutations, as sequencing confirmed the original deletion remained and no compensatory mutations arose in Gag. Expectedly, the mutants showed around a 50% reduction in packaging and a reduction in the ability to form discrete dimers on Northern blots, which the authors hypothesised to be due to packaging of spliced RNA within the virion.

The authors suggest that a cellular factor may bind to SL1 to enhance HIV-1 replication, although further exploration of this is complicated by the fact that SL1 sits at the heart of a complex RNA structure regulating multiple aspects of the virus's lifecycle.

In addition, DIS mutation may disrupt the balance between the two structural conformations of the 5'UTR (Figure 6). This may weaken the interaction between the DIS and the U5 helix, (Kenyon et al., 2013; Lu et al., 2011), thus promoting dimerisation by allowing SL1 to be exposed for inter-strand base pairing.

The effect of Stem B and Loop B mutations on dimerisation has not been assessed in 293T cells before, and the results (Figure 17) show that they exhibited dimerisation defects as severe as those previously observed in COS7 cells (Table 9) (Shen et al., 2001). As in Hill et al (Hill et al., 2003), these experiments suggest that the effect of particular SL1 mutations (in this instance the DIS mutation) can be cell-type dependent. However, the Stem B and Loop B mutants exhibited similar phenotypes to those observed

in COS7 cells, suggesting that these mutations play a more fundamental role in dimerisation that is not cell type-dependent.

Table 9 – comparison of dimerisation and packaging data from 293T and COS7 cells

Mutant	Cell type	Dimerisation (% wild type)	Packaging (% wild type)	Reference
DIS	293T	100% 110%	- 104%	(Song et al., 2007) This study
	COS7	59% -	78% 60%	(Shen et al., 2000) (Laughrea et al., 1999)
Stem B	293T	48%	34%	This study
	COS7	59%	25%	(Shen et al., 2001)
Loop B	293T	77%	52%	This study
	COS7	59%	44%	(Shen et al., 2001)

Dimerisation and packaging efficiencies from this study and others are expressed relative to wild type.

3.4.5. Effect of compensatory mutations

Two compensatory mutations that arose during long term culture of SL1 mutants and restored infectivity (Liang et al., 1998) but not dimerisation (Shen et al., 2000) were introduced into each of the SL1 mutants and the wild type virus. CA-SP1 processing defects in the Loop B and LD4 mutants were corrected (Figure 18) but processing was also accelerated in the other mutants and the wild type, showing that they act whether SL1 is completely intact, subtly or grossly mutated.

Correction of defective processing was associated with restoration of infectivity (Figure 19) suggesting that delayed processing is a major contributor to the attenuation of these mutants. No change was observed in the infectivity of the DIS mutant and only a small increase in the Stem B mutant, possibly related to the ability of the mutations to increase packaging efficiency (Liang et al., 2000, 1998; Shen et al., 2000).

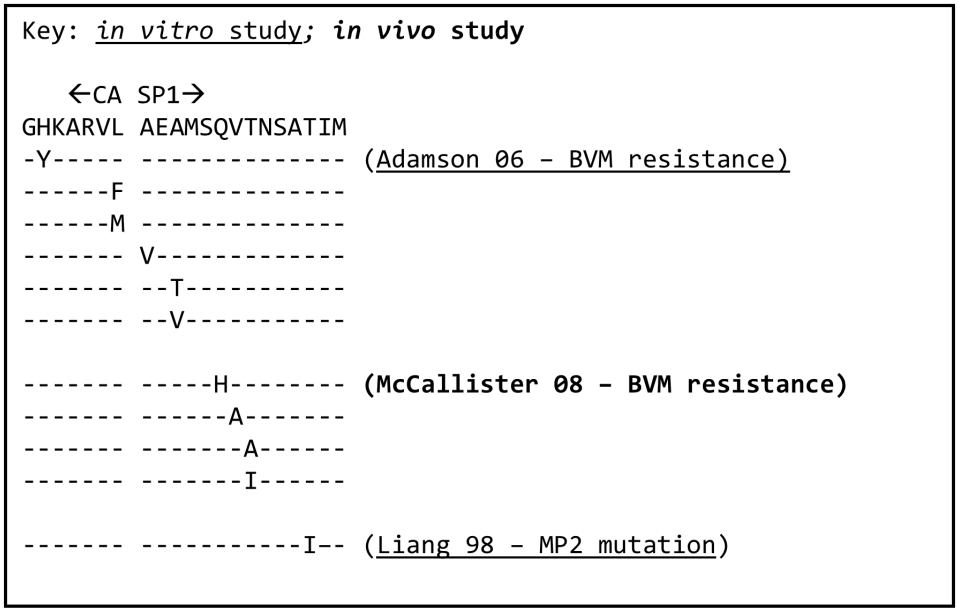
It is intriguing that the introduction of mutations which accelerate CA-SP1 cleavage had no effect on wild type virus infectivity (Figure 19) since infectivity is tightly linked to optimal Gag processing (Li et al., 2003; Nowicka-Sans et al., 2016; Wiegers et al., 1998). This is investigated further in the next chapter.

3.4.6. Parallels between SL1 compensatory mutations and maturation inhibitor resistance mutations

The effect of the Loop B and LD4 mutations on CA-SP1 processing raises striking parallels with the effects of MIs, which attenuate virus infectivity by specifically blocking this cleavage step. Furthermore, the MP2 compensatory mutation that is involved in correcting defective processing maps to the same region of Gag as resistance mutations to the prototype MI bevirimat (Figure 22) (Adamson et al., 2006; Li et al., 2003; McCallister et al., 2008) and, like MP2, some of these result in specific acceleration of CA-SP1 processing (Adamson et al., 2006; Li et al., 2003). Resistance mutations also emerged following

long-term culture of infected cells with another MI, PF-46396, including one in CA_{CTD} and one in SP1 (Waki et al., 2012).

Figure 22 – Location of bevirimat resistance mutations and SL1 compensatory mutation MP2



Sequence of CA_{CTD} and SP1, highlighting resistance mutations which emerged following in vitro culture (Adamson et al., 2006) and in vivo treatment (McCallister et al., 2008) with bevirimat (BVM). For comparison the position of the MP2 mutation (Liang et al., 1998) is indicated.

The precise mechanism of action of bevirimat is unknown, but analysis of binding sites suggests that it blocks CA-SP1 processing by stabilising the CA-SP1 junction α-helix preventing exposure of the cleavage site (Schur et al., 2016; Wagner et al., 2016). Resistance mutations map to CA-SP1 protein interfaces rather than drug binding sites, and are hypothesised to act by destabilising the helix at the CA-SP1 junction, to cancel out the helix-stabilising effects of the drug (Schur et al., 2016).

Given the phenotypic effects of SL1 mutation and the location of compensatory mutations, one attractive possibility is that the Loop B and LD4 mutations result in over-stabilisation of the CA-SP1 helix, and that the MP2 mutation promotes destabilisation.

4. Effect of SL1 mutation on ESCRT function

4.1. Introduction

4.1.1. Parallels between mutation of SL1 and late domain motifs

CA-SP1 processing defects are regarded as hallmarks of late domain mutants, as they accompany the defect in particle release caused by p6 mutation (Fisher et al., 2007; Göttlinger et al., 1991; Huang et al., 1995). It is striking that the Loop B and LD4 mutants also exhibit this phenotype (Figure 15), which raises an interesting possibility that SL1 may interact with the ESCRT pathway directly or indirectly and be implicated in virus budding.

4.1.2. Late domain mutant-like phenotype of NC mutants

NC mutations also give rise to a late domain mutant-like phenotype (Dussupt et al., 2011). When viewed under the electron microscope virus particles have an immature morphology and are observed tethered to the membrane indicating a failure to sever budding necks.

There is a positive correlation between the number of basic NC residues mutated, the packaging of gRNA and the severity of Gag processing defects (Cimarelli et al., 2000). Replacement of single residues in NC with alanine causes delayed processing of MA-CA-SP1 and CA-SP1, as well as a modest two-fold reduction in virus release (Dorfman et al., 1993). Replacing the distal cysteine residues of both zinc fingers results in a 10-fold reduction in release from HeLa cells (Dorfman et al., 1993) although the effect is more modest in 293T cells (Popov et al., 2008). Replacement of all basic residues with alanine in the zinc fingers of NC results in a 13-fold reduction in virus release from 293T and HeLa cells. In Jurkat and CEM T cells NC mutant virus release was completely blocked (Dussupt et al., 2011). Given that NC's main role is in RNA binding it is tempting to speculate that this function is important for ESCRT recruitment.

These observations are interesting in the context of this study as SL1 is an integral part of the structure recognised by NC for packaging (Section 1.12) (Kim et al., 1994; Sakaguchi et al., 1993). Thus, it appears that mutation of *cis*- or *trans*- acting packaging signals can disturb Gag processing.

4.1.3. NC interacts with ALIX and TSG101

NC appears to be involved in recruitment of ALIX and TSG101 by Gag, as both ALIX (Popov et al., 2008) and TSG101 (Chamontin et al., 2015) physically interact with NC in *in vitro* pulldown assays. Little is known about the latter interaction, but the ALIX-NC interaction is becoming increasingly well understood.

ALIX is incorporated into YPX_nL^- mutant virions, providing the zinc finger motifs are intact, demonstrating that ALIX interacts with Gag through the NC domain in addition to the YPX_nL motif (Dussupt et al., 2009). When both the PTAP and YPX_nL motifs are mutated, budding can be restored by

expressing the Bro1 domain of ALIX alone, providing NC is intact, as Bro1 contains a binding site for the ESCRT-III protein CHMP4 (Dussupt et al., 2009).

Although both TSG101 and ALIX may interact with NC, the NC-ALIX interaction appears to be stronger than that of NC-TSG101. This is because mutation of the YPX_nL motif while maintaining NC only moderately reduces the amount of ALIX incorporated into virions, whereas mutation of the PTAP motif almost completely abolishes TSG101 incorporation (Bendjennat and Saffarian, 2016). However deletion of the second zinc finger of NC severely reduced colocalization of Gag and TSG101 at the plasma membrane (Chamontin et al., 2015).

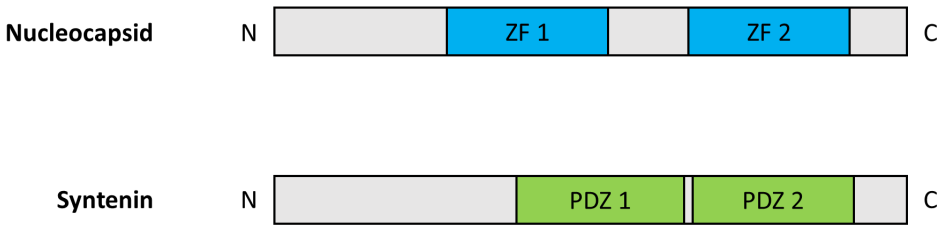
4.1.4. Possible role of RNA in ALIX/TSG101-Gag interaction

The interaction between ALIX and NC was initially thought to be direct because it was not abolished by nuclease treatment (Popov et al., 2008). However, solvent accessible surface analysis combined with mutagenesis revealed 5 residues important for NC binding forming a positively charged patch in the N-terminal domain of Bro1. As the NC residues responsible for the interaction are also positively charged, it seemed plausible that a negatively charged intermediate such as nucleic acid could bridge these two proteins. Indeed, in contrast with the previous study RNase treatment was found to disrupt the interaction (Sette et al., 2012).

This discrepancy was clarified later when it was found that the NC-Bro1 interaction can be mediated by RNA or lipids (Sette et al., 2016). In the first study (Popov et al., 2008) lipids were present in the sample preparation rendering the interaction insensitive to RNase treatment, whereas in the second study (Sette et al., 2012) lipids were excluded by detergent treatment. Nuclease treatment enhanced binding of NC to membrane lipids (Sette et al., 2016), in a manner analogous to the negative regulation of MA binding to internal membranes by tRNA binding (Chukkapalli et al., 2010; Kutluay et al., 2014).

Using a yeast two-hybrid screen Sette *et al* found that ALIX's natural binding partner is Syntenin, which NC appears to mimic (Sette et al., 2016). Syntenin is involved in cell extension, interacting through its tandem PDZ motifs, analogous to the zinc fingers of NC (Figure 23), with the Bro1 domain of ALIX in the plasma membrane.

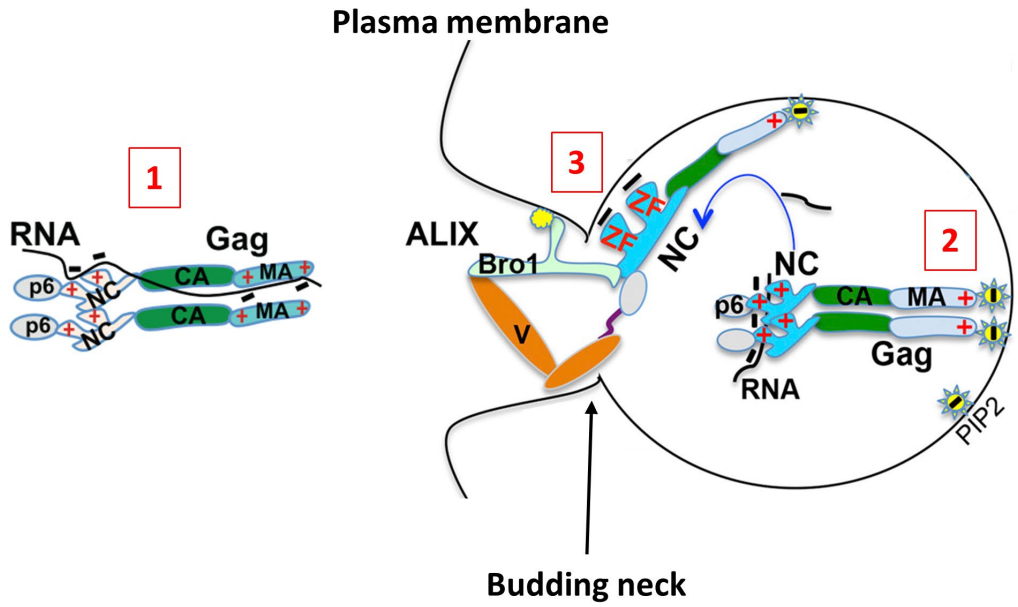
Figure 23 – Comparison of nucleocapsid and Syntenin



Cartoon representation of nucleocapsid and Syntenin with domains involved in the Bro1 interaction coloured. Nucleocapsid and Syntenin both contain tandem interaction domains. Nucleocapsid interacts with Bro1 through its zinc finger motifs, while Syntenin uses PDZ domains.

The authors proposed a model whereby a subset of Gag molecules in the assembled immature virion switch binding preference from RNA to plasma membrane lipids at the budding neck, enabling recruitment of ALIX and subsequent binding of ESCRT-III proteins to facilitate scission (Figure 24). However, this model does not account for the finding that RNA has been shown to act as a bridge between ALIX and NC *in vitro*. We propose that RNA binding may maintain NC in a conformation which is suitable for Bro1 binding, before exchange of an RNA bridge for a lipid one in the budding neck.

Figure 24 – Model of ALIX function in HIV-1 budding proposed by the Bouamr lab



1) In the cytoplasm, Gag assembly complexes associate with RNA, stimulating Gag nucleation and preventing premature association with internal membranes. 2) When Gag reaches the plasma membrane, it becomes enriched in PI(4,5)P2 domains. 3) Once sufficient Gag has accumulated to form a spherical particle, a subset of NC domains find themselves in close proximity to the budding neck, where they dynamically trade RNA for lipids at the plasma membrane. NC and the Bro1 domain of ALIX co-insert in these cholesterol rich domains, stimulating ESCRT-III recruitment and membrane scission. Image modified from (Sette et al., 2016).

4.2. Experimental aims

Taken together, the studies discussed above show that RNA can bridge the ALIX-NC interaction *in vitro*, and raise the possibility that RNA is involved in enabling Gag to recruit ALIX *in vivo*. The CA-SP1 processing defect observed in the Loop B and LD4 mutants (Figure 15) is reminiscent of that observed in late domain and NC mutants, so it was of interest to investigate specifically if SL1 mutation affects virus release and the function of ALIX. To assess the effect of SL1 mutation on virus release, Western blotting was performed. The Gag processing phenotypes of SL1, NC and late domain mutants were compared, and the effect of combining the LD4 mutation with a PTAP late domain mutation on Gag processing was also studied to determine if SL1 and late domain mutations disrupt the same pathway. An ALIX overexpression assay was used to determine the importance of SL1 in ALIX-mediated virus release and enhancement of CA-SP1 processing. Finally, the compensatory mutations were introduced in the context of mutated late domains and knocked-down ESCRT proteins to assess the requirement of ESCRT for their ability to rescue SL1 mutant phenotypes.

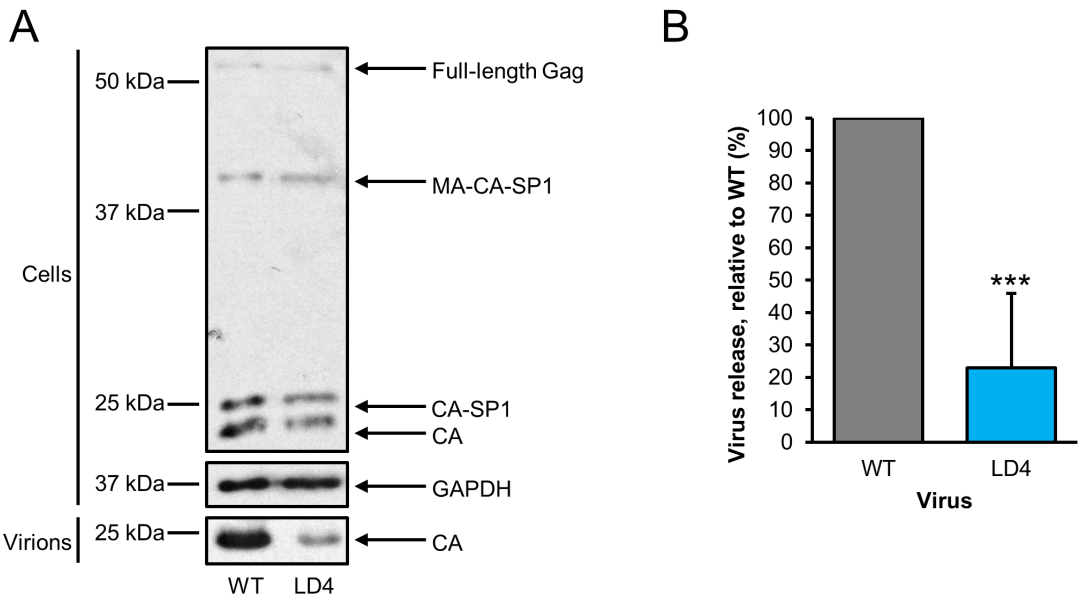
4.3. Results

4.3.1. The LD4 mutation reduces virus release

For these experiments the LD4 mutant was used as it exhibited the greatest defect in CA-SP1 processing (Figure 15). Densitometry was performed to quantify the release ratio – the amount of extracellular Gag relative to intracellular Gag – from Western blots of cell and virion-associated Gag (Figure 25).

The release ratio of the LD4 mutant was found to be about 4-fold lower than wild type (Figure 25). However, the reduction in extracellular pelletable CA may be due to reduced virion stability, as observed for some NC mutants (Wang and Aldovini, 2002). To exclude this possibility an assay based on a previously published protocol was used (Ohagen and Gabuzda, 2000). Briefly, purified virions were treated with detergent to break down membranes, followed by centrifugation to pellet intact cores, leaving CA from degraded virions in the supernatant. Virion stability was assessed by calculating the proportion of total pelleted CA (**Figure** 26).

Figure 25 – The LD4 mutation in SL1 reduces virus release from 293T cells



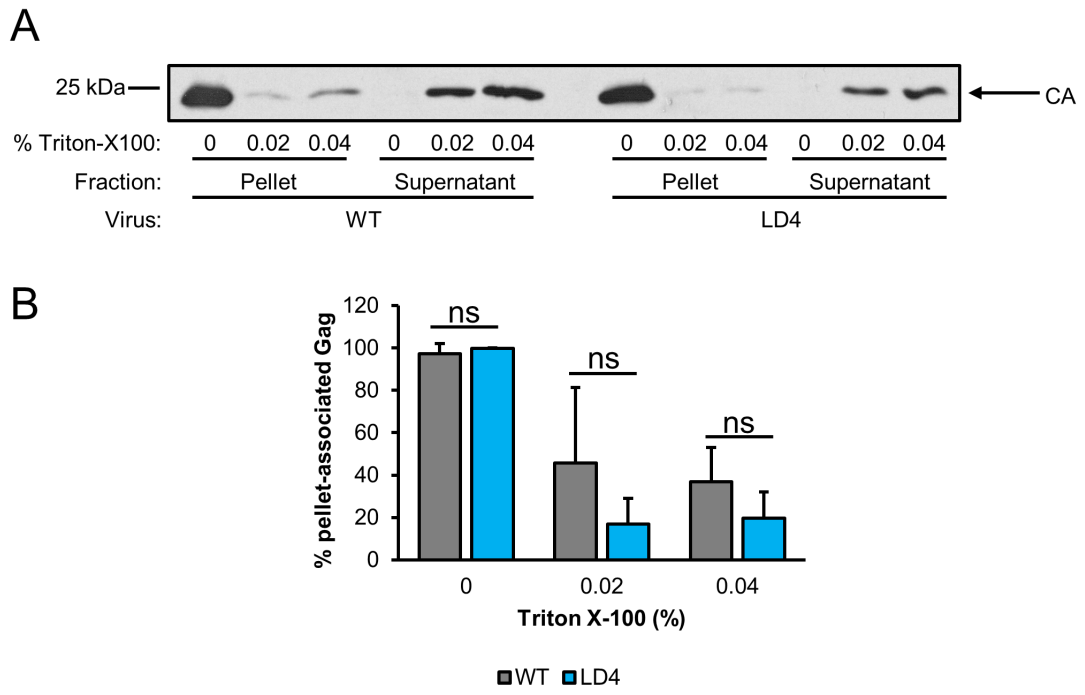
A) Western blot of cell lysates and purified virions following transfection with the indicated proviruses. B) Quantitation of release ratios. Densitometry was performed on all Gag associated bands and release ratios were normalised to wild type. The graph shows the mean values from 5 independent experiments. Error bars represent standard deviation.

Virions were not degraded in the absence of detergent treatment, in agreement with previous studies (Ohagen and Gabuzda, 2000; Willey et al., 1994). Detergent treatment caused appearance of CA in the supernatants, indicating dissociation of viral cores following membrane disruption. On average wild type virus preparations pelleted two-fold more CA than the LD4 mutant. This difference was not

statistically significant, however with further repeats of the experiment this may not be the case. Taken together, results from Figure 25 and Figure 26 suggest that the reduction in CA pelleted by the LD4 mutant is likely to be due to a genuine reduction in virus release rather than reduced virion stability, although a small effect on capsid stability cannot be excluded given the trends observed in the data.

SL1 mutation therefore gives rise to defects in Gag processing and a reduction in virus release that are both typically associated with late domain mutations.

Figure 26 – Stability of wild type and LD4 capsids and release into supernatants

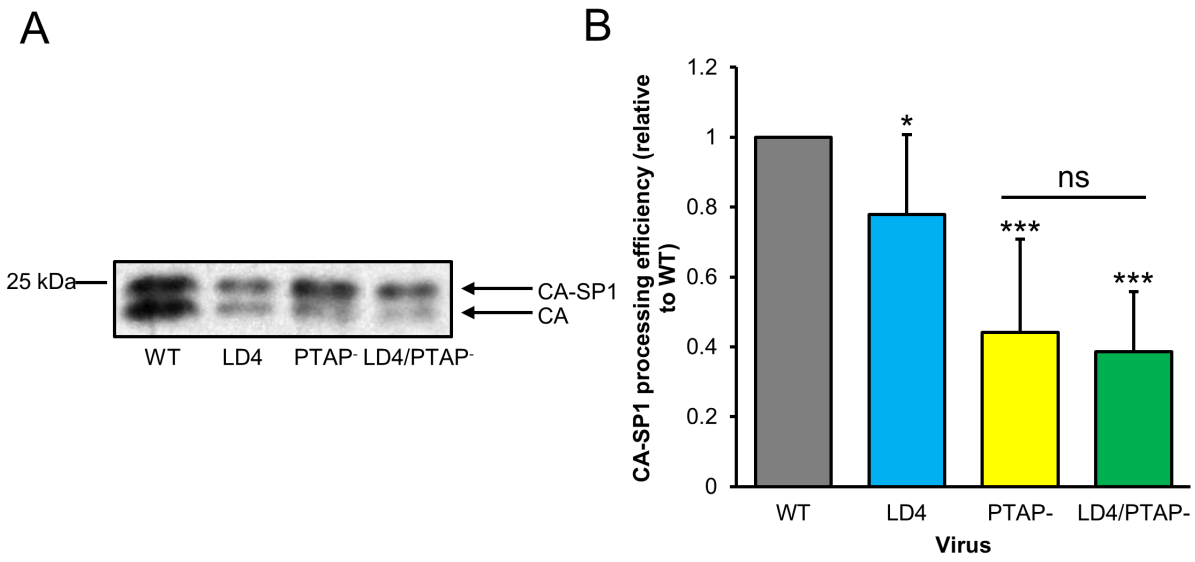


A) Western blot showing pellet and supernatant-associated CA bands for each mutant following treatment with the indicated concentrations of detergent. B) Quantitation of the proportion of CA pelleted for each mutant under each condition. The graph shows the mean values from 3 independent experiments. Error bars represent standard deviation.

4.3.2. SL1 and late domain mutation effects are non-additive

It was noted previously that combining NC and PTAP mutations did not have an additive effect on the defect in processing, indicating that NC and p6 domain mutations disrupt the same pathway (Popov et al., 2008). To investigate if the same is true for SL1 and NC mutations, CA-SP1 processing efficiency was calculated for an LD4/PTAP$^-$ double mutant (Figure 27).

Figure 27 – Non-additive effect of LD4 and PTAP$^-$ mutations on CA-SP1 processing



A) Western blot of cell lysates following transfection with the indicated proviruses. B) Quantitation of CA-SP1 processing efficiency normalised to wild type. The graph shows the mean values from 9 independent experiments. Error bars represent standard deviations.

When combined, the processing defect of the double mutant was almost identical to the one observed for the PTAP$^-$ mutation alone (Figure 27).

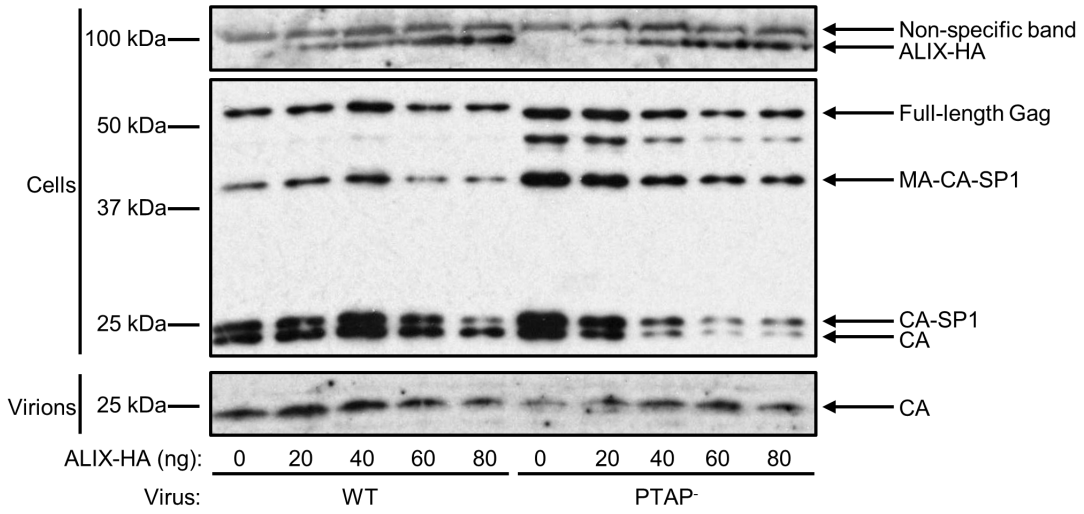
4.3.3. Mutation of SL1 reduces ALIX-dependent virus release

As ALIX has been shown to bind to NC in an RNA-dependent manner (Sette et al., 2012), it was hypothesised that an intact SL1 is required for efficient ALIX binding to NC.

To test this hypothesis, an ALIX overexpression assay was used. Mutation of the PTAP motif blocks the TSG101 pathway making the virus dependent on overexpression of ALIX for otherwise inefficient virus release. By combining a PTAP$^-$ mutation with other mutations it is possible to identify sequences required for efficient ALIX function. This approach was previously used to discover that NC contains an ALIX binding site (Popov et al., 2008).

First, the assay was optimised by transfecting increasing amounts of HA-tagged ALIX-expressing plasmid alongside wild type and PTAP$^-$ provirus plasmids (Figure 28).

Figure 28 – Optimising rescue of release by ALIX overexpression

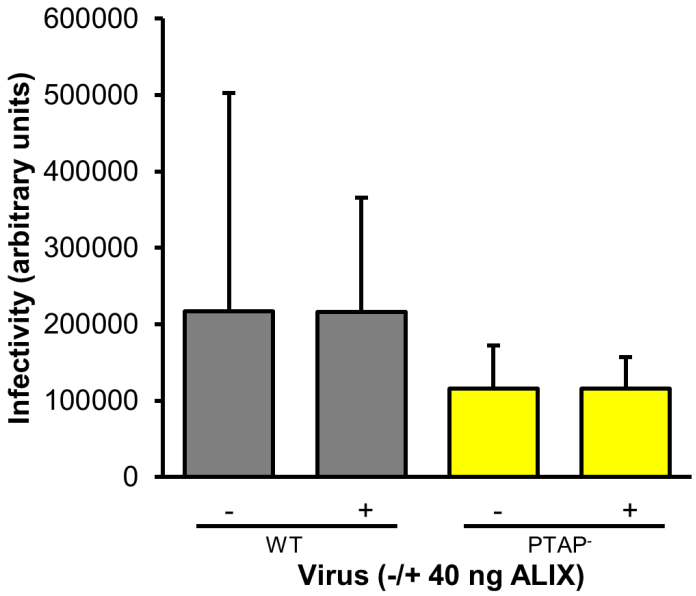


Western blot of cell lysates and purified supernatants, following transfection with the indicated proviruses and increasing amounts of ALIX-HA plasmid.

ALIX overexpression had no effect on wild type virus release, however at transfection amounts greater than 60 ng there was an inhibitory effect as previously reported (Dussupt et al., 2009). PTAP$^-$ mutant release was most efficient when 40-60 ng of ALIX-HA was transfected, so for the overexpression assay 40 ng of ALIX-HA plasmid was used.

Despite rescuing virus release (Figure 28), overexpression of ALIX appeared to have no effect on virus infectivity (Figure 29), however there was a great deal of variation particularly for the WT virus without ALIX overexpression (bar 1 in Figure 29).

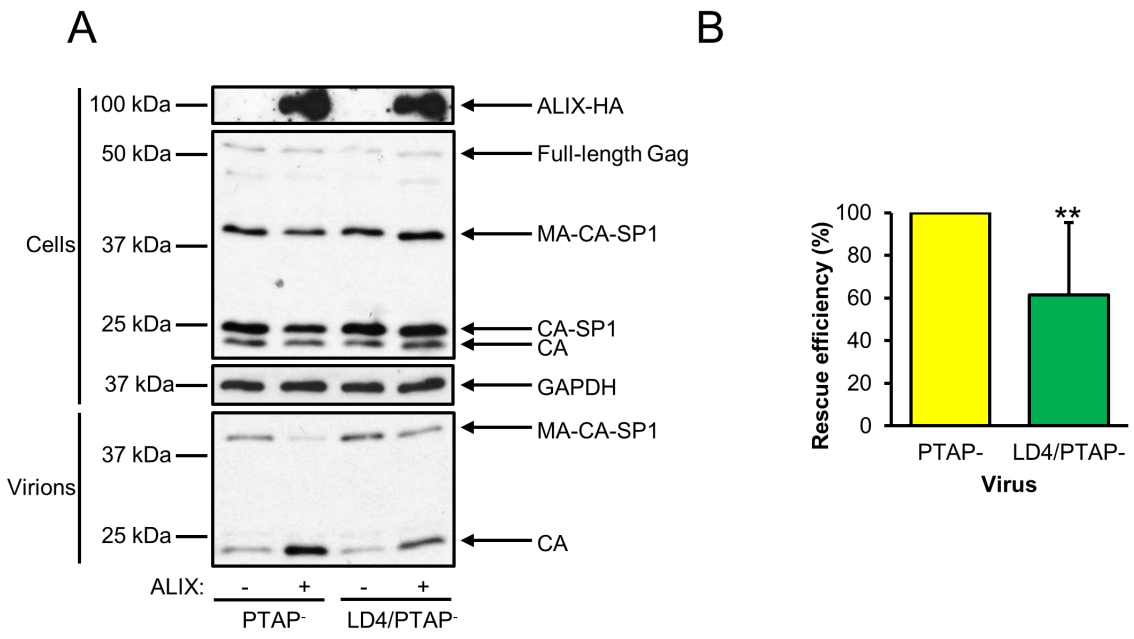
Figure 29 – Effect of ALIX overexpression on virus infectivity



Infectivity of the wild type virus and PTAP$^-$ mutant produced by transfection of 293Ts with provirus alone (bars 1 and 3) or cotransfected with 40 ng ALIX plasmid (bars 2 and 4). The graph shows the mean values from 3 independent experiment (2 experiments for bar 4). Error bars represent standard deviation.

The ability of overexpressed ALIX to rescue release of PTAP$^-$ and LD4/PTAP$^-$ mutants was then assessed. The fold change in release ratio upon ALIX overexpression was calculated and normalised to the PTAP$^-$ control (Figure 30).

Figure 30 – Mutation of SL1 reduces the ability of ALIX to rescue release of the PTAP$^-$ mutant



A) Western blot of cell lysates and purified supernatants following transfection with the indicated proviruses, with or without ALIX overexpression. B) Quantitation of the efficiency of rescue by ALIX. Densitometry was performed on all p24-associated bands from lysates and virions. The graph shows the mean values from 9 independent experiments. Error bars represent standard deviation.

Introduction of the LD4 mutation did not block ALIX's ability to rescue release, however, there was a modest but significant reduction in the efficiency of rescue (Figure 30). The band with a molecular weight equivalent to that of MA-CA-SP1 on the lower virion blot is unusual, as virion preparations normally only contain fully processed Gag in the form of CA. It is possible that it is a cellular artefact but given its similarity in size to MA-CA-SP1 it has been retained in the figure. Nevertheless, this band does not affect the interpretation that the ability of ALIX to rescue virus release is weaker in the presence of the LD4 mutation.

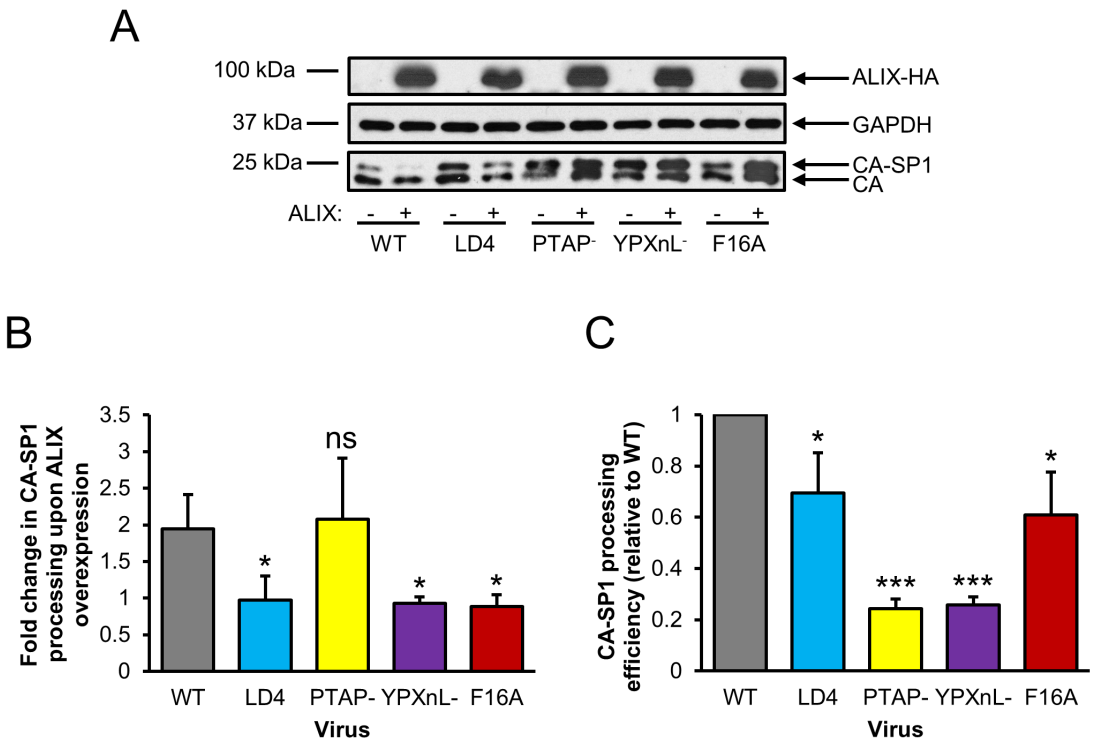
In agreement with Figure 27, combining the PTAP$^-$ mutation with the LD4 mutation (lane 3 of Figure 30) did not worsen the CA-SP1 processing defect of the PTAP$^-$ mutant. The lower panel of the blot in Figure 30 shows that the effect of combining the mutation on virus release is also not additive, as similar amounts of Gag are present in the supernatant of the PTAP$^-$ and LD4/PTAP$^-$ mutants.

4.3.4. The LD4 processing defect is comparable to that of NC mutation and neither can be rescued by ALIX overexpression

In performing ALIX overexpression experiments (Figure 28 and Figure 30) it was found that overexpression of ALIX accelerated CA-SP1 processing. Measuring the ability of ALIX to accelerate processing may be another way to determine which sequences in the viral genome are required for its function. Accordingly, cells were transfected with proviruses containing mutations in ALIX binding sites

(YPX$_n$L$^-$ in p6 and the first zinc finger of NC (F16A)), the PTAP motif (serving as a negative control), and the LD4 mutant (Figure 31).

Figure 31 – Requirements for acceleration of CA-SP1 processing by ALIX



A) Western blot of cell lysates following transfection with the indicated proviruses, with or without ALIX-HA. B) Quantitation of the fold-change in CA-SP1 processing upon overexpression of ALIX. The graph shows the mean values from 3 independent experiments. Error bars represent standard deviation. C) Quantitation of the efficiency of CA-SP1 processing relative to WT. The graph shows the mean values from 3 independent experiments. Error bars represent standard deviation.

ALIX overexpression increased the efficiency of CA-SP1 processing by two-fold in the wild type virus and the PTAP$^-$ mutant (Figure 31**B**). However, mutation of either NC or the YPX$_n$L motif prevented ALIX from accelerating processing, as did mutation of SL1.

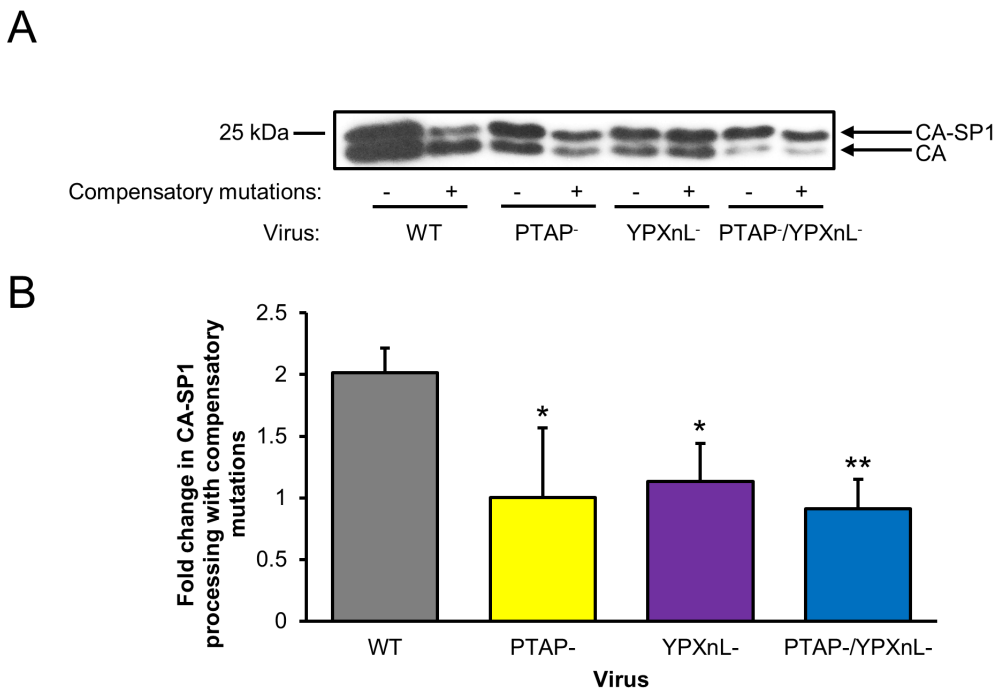
In the absence of ALIX overexpression, the PTAP$^-$ and YPX$_n$L$^-$ mutations resulted in the most severe CA-SP1 processing defects, whereas the LD4 mutant possessed a much milder CA-SP1 processing defect of comparable intensity to the F16A NC mutant (Figure 31**C**).

4.3.5. Compensatory mutations require intact late domains and TSG101 and ALIX to restore CA-SP1 processing

Compensatory mutations emerging from SL1 mutant cultures (Liang et al., 2000, 1998) rescue defective virus replication (Figure 19) and restore CA-SP1 processing (Figure 18). The data shown above suggest that SL1 may interact with the ESCRT pathway. It was of interest therefore to investigate if the

compensatory mutations cooperate with the ESCRT pathway by evaluating their ability to accelerate processing in the presence of mutated late domain motifs (Figure 32).

Figure 32 – Late domain-dependence of compensatory mutations

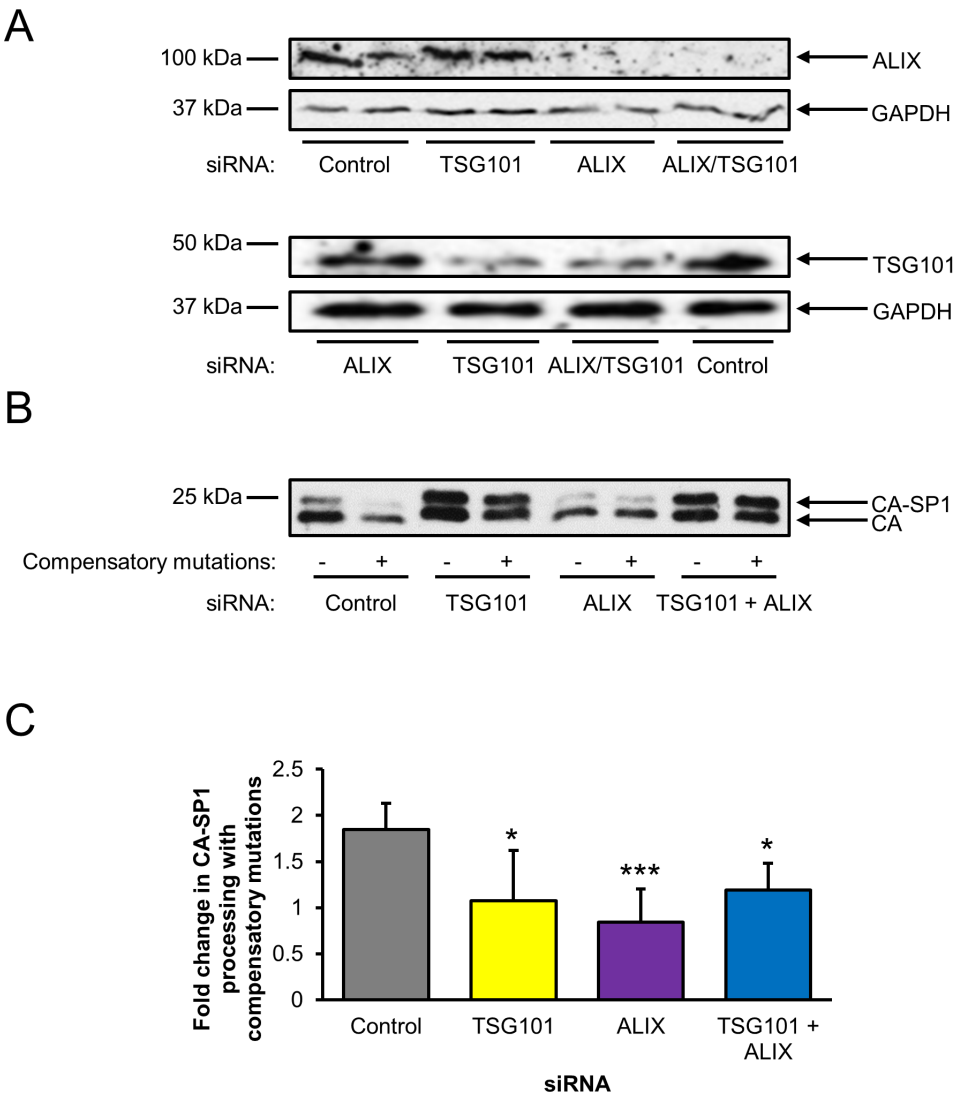


A) Western blot of cell lysates following transfection with the indicated proviruses. B) Quantitation of the fold-change in CA-SP1 processing upon introduction of compensatory mutations. The graph shows the mean values from 3 independent experiments. Error bars represent standard deviation.

Unlike in the wild type virus where acceleration of CA-SP1 processing was observed, the compensatory mutations were unable to correct the processing defects observed in the single or double late domain mutants (Figure 32).

As the PTAP and YPX$_n$L motifs interact with TSG101 and ALIX respectively, siRNA transfections were performed to knock down these proteins in the presence or absence of compensatory mutations (Figure 33B & C). To test knockdown efficiency cells were transfected with the various siRNAs and Western blotting was performed with anti-ALIX and anti-TSG101 antibodies (Figure 33A).

Figure 33 – ESCRT-dependence of compensatory mutations

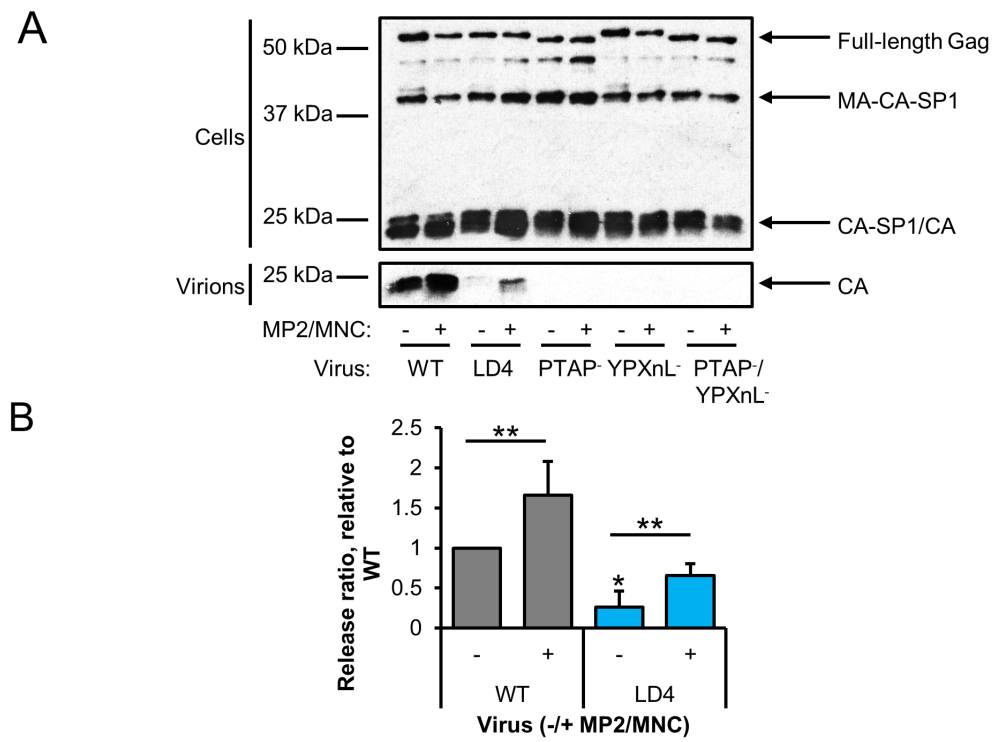


A) Western blot of cell lysates following transfection with the indicated proviruses and siRNAs. B) Quantitation was performed as in Figure 32. The graph shows the mean values from 3 independent experiments. Error bars represent standard deviation.

Consistent with the results in Figure 32, knockdown of TSG101 and/or ALIX resulted in a failure to accelerate processing. Differences were observed in the intensity of Gag bands following different siRNA treatments. Knockdown of TSG101 potently inhibits virus release resulting in intracellular accumulation of Gag (Garrus et al., 2001), whilst knockdown of ALIX has been previously observed to result in lower levels of intracellular Gag likely due to reduced cell viability (Martin-Serrano et al., 2003).

The compensatory mutations were also tested for their ability to rescue the release of the late domain mutants, in addition to the wild type and LD4 mutant (Figure 34).

Figure 34 – Effect of compensatory mutations on virus release



A) Western blot of cell lysates and purified virions following transfection with the indicated proviruses. B) Quantitation of the release of wild type and LD4 mutant virion, with and without compensatory mutations. Densitometry was performed on all Gag bands from virion and lysate Westerns. The graph shows the mean values from 6 independent experiments. Error bars represent standard deviation.

Consistent with failure to accelerate processing, the compensatory mutations did not rescue release of late domain mutant viruses. They did however partially rescue the release of the LD4 mutant (Figure 34), and surprisingly they also increased release of wild type virus.

4.4. Discussion

4.4.1. SL1 mutants have a late domain mutant-like phenotype

As CA-SP1 processing defects were observed in SL1 mutants (Figure 15), the effect of SL1 mutation on virus release was investigated. There was approximately a 4-fold reduction in the release ratio of the LD4 mutant compared to wild type (Figure 25). This effect was not due to particle instability (Figure 26). It was hypothesised that the late domain mutant-like phenotype caused by SL1 mutation may be due to the disruption of the ESCRT pathway.

To investigate this, a mutant provirus containing both the LD4 mutation and the PTAP$^-$ mutation was created. Combining the mutations did not produce an additive effect on CA-SP1 processing (Figure 27), suggesting that they do indeed act through a shared mechanism to delay processing.

4.4.2. SL1 is involved in efficient ALIX function

One plausible mechanism by which SL1 might be involved in the ESCRT pathway is through interactions with the positively charged interfaces in NC and the Bro1 domain of ALIX. RNA has been shown to be capable of mediating the ALIX-NC interaction *in vitro* (Sette et al., 2012) although the relevance of an RNA-mediated interaction *in vivo* is not known.

To explore this an ALIX rescue assay was performed. The rationale behind this experiment was that if SL1 is required for ALIX to function, ALIX would be less able to rescue release of a mutant containing an SL1 mutation.

The data show that release of both mutants could be increased upon expression of ALIX-HA, but the LD4/PTAP$^-$ mutant was rescued with an average efficiency of 60% relative to the PTAP$^-$ mutant (Figure 30). However, this rescue was not completely blocked suggesting that an intact SL1 is not essential for ALIX function, but that it is required for efficient function.

This contrasts with data showing that mutation of the NC zinc fingers severely reduces the ability of ALIX to rescue release (Popov et al., 2008) most likely by blocking the lipid-mediated interaction hypothesised to occur between NC and ALIX for recruitment of ESCRT-III to the budding neck (Figure 24) (Sette et al., 2016). As the NC domain of the LD4 mutant is not disrupted, this would presumably not affect the lipid-mediated interaction, explaining the discrepancy between the NC and SL1 data. However, the observation that SL1 mutation does impair rescue suggests that RNA may be important for the interaction between NC and ALIX at the late stage of virus release.

4.4.3. SL1 mutation produces a similar phenotype to NC mutation

Mutation of NC also results in an accumulation of CA-SP1 and reduced virus release (Dorfman et al., 1993). SL1 is one of two stem-loops important for packaging, so the observation that mutation of either *cis*- and *trans*- acting packaging signals results in similar phenotypes suggests that the defects observed may be caused by disruption of the interaction between NC and psi.

To compare the effects of various mutations on the severity of the CA-SP1 processing defect, NC mutant F16A as well as the LD4, PTAP$^-$ and YPX$_n$L$^-$ p6 mutants were assessed. The F16A mutation is in the first zinc finger motif and reduces gRNA binding (Dorfman et al., 1993) but has no effect on non-specific RNA binding (Cimarelli et al., 2000). The LD4 mutation causes a milder CA-SP1 processing defect than classic late domain mutations, resembling the defect causes by the F16A mutant (Figure 31). The Bro1 domain of ALIX interacts with NC through its zinc fingers (Popov et al., 2008), and as the residue substituted in the F16A mutant is within zinc finger 1 and is critical for its function (Dorfman et al., 1993) it is tempting to speculate that the defect in CA-SP1 processing caused by the F16A mutation is also due to disruption of this interaction. A possible explanation for the milder phenotype

of the LD4 and F16A mutants is that both contain a wild type p6 domain, which is critical for recruiting ESCRT components for virus release.

4.4.4. Motifs required for ALIX to accelerate CA-SP1 processing

Overexpression of ALIX accelerated CA-SP1 processing (Figure 31), possibly due to enhanced recruitment of ESCRT-III (Morita et al., 2011). ALIX was overexpressed in different contexts and the effect on processing observed to determine which motifs are required for its function.

On average, ALIX overexpression caused a 2-fold increase in CA-SP1 processing in the wild type and $PTAP^-$ mutant (Figure 31). This was not observed in the YPX_nL^- mutant. The F16A mutation in NC also blocked acceleration of processing, suggesting that specific binding of gRNA to NC is required for efficient ALIX function. Likewise, acceleration was also blocked in the LD4 mutant, supporting the finding that the LD4 mutation reduced ALIX's ability to rescue release (Figure 30).

4.4.5. ALIX and TSG101 are required for rescue of late domain mutant-like phenotype by compensatory mutations

It was hypothesised that if SL1 mutation disrupts ESCRT function, then the MP2 and MNC compensatory mutations which restore Gag processing and infectivity to SL1 mutants might act through a mechanism involving ESCRT. The compensatory mutations were introduced into late domain mutants to see if they were capable of correcting the defective Gag processing associated with these mutants.

Interestingly, the compensatory mutations failed to enhance CA-SP1 processing when the PTAP and YPX_nL motifs were mutated individually or in combination (Figure 32) or when the motifs' binding partners TSG101 and ALIX were knocked down (Figure 33). They also failed to rescue the release of late domain mutants, but they increased release of both the wild type and LD4 mutant viruses (Figure 34). This suggests that binding of ALIX and TSG101 to the p6 domain of Gag is required for the compensatory mutations to be able to correct the defective Gag processing and budding caused by SL1 mutation.

The apparent requirement for ALIX and TSG101 to rescue SL1 mutants may be due to compensatory mutations restoring an interaction with these proteins which is lost upon SL1 mutation. It is known that the MNC mutation enhances packaging efficiency (Liang et al., 2000, 1998; Shen et al., 2000) and the MP2 mutation enhances packaging specificity (Roy et al., 2006; Russell et al., 2003), so restoration of gRNA binding may restore ESCRT binding to NC.

As ALIX can bind to NC in an RNA-dependent manner (Chamontin et al., 2015; Popov et al., 2008; Sette et al., 2012) the compensatory mutations may indirectly restore the interaction between these proteins enabling increased recruitment of ESCRT-III and restored Gag processing and virus release.

These data provide further evidence of a link between the packaging signal in the 5'UTR and the p6 domain of Gag, as they show that mutations in SL1 are rescued by the evolution of mutations which act through a mechanism requiring the p6 domain and its host cell protein binding partners.

It is intriguing that compensatory mutations which disrupt wild type Gag processing by accelerating it have no effect on virus infectivity (Figure 19). Interference with the coordinated process of Gag processing is usually detrimental to infectivity (Li et al., 2003; Nowicka-Sans et al., 2016; Wiegers et al., 1998). Data from a recent study show that delaying budding causes critical virion components such as RT and IN to escape the virion before the budding neck closes (Bendjennat and Saffarian, 2016). If the rate of processing alone was accelerated, the result might be the same. However, if increased processing is matched by increased budding, the two processes would be in equilibrium enabling the virus to remain fully infectious as shown in Figure 19.

5. Effect of gRNA and 5'UTR expression in trans on Gag processing and virus release

5.1. Introduction

The data presented in the previous chapters suggest that correct binding of NC to SL1 is required for efficient ESCRT recruitment, stimulating virus maturation and release. Expression of Gag in the absence of gRNA is sufficient for formation and budding of VLPs from cells (Gheysen et al., 1989). However, following cotransfection of labelled Gag and gRNA, 90% of released virions were observed to contain gRNA (Chen et al., 2009), indicating that there is a strong preference for release of virions containing successfully packaged viral genomes. Coordination of packaging with ESCRT recruitment may serve as a quality control mechanism by preferentially releasing gRNA-containing virions, maximising the opportunities for productive infection of new host cells whilst avoiding unnecessary activation of the immune system by empty virions.

If a Gag-gRNA interaction is important for ESCRT recruitment, then one might predict that artificially increasing the levels of gRNA in the cytoplasm would influence CA-SP1 processing and virus release. In support of this, a previous study found that Gag-expressing cells released more VLPs when cotransfected with a construct expressing a truncated version of the HIV-1 genome (Ueno et al., 2004). However, the effect on Gag processing was not investigated as the Gag construct used lacked PR.

5.2. Experimental aims

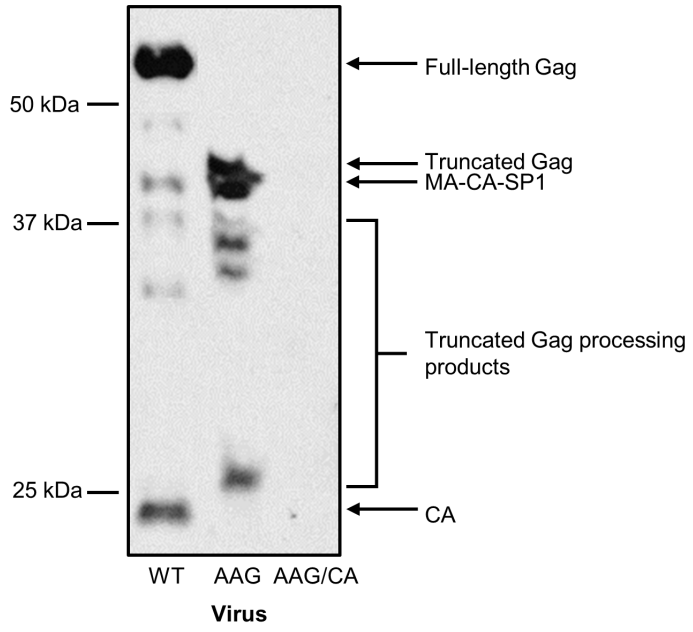
To investigate the effect of *in trans* overexpression of gRNA on Gag processing and virus release, cells were transfected with proviruses and a construct expressing full length gRNA containing mutations to prevent Gag translation. Western blotting was performed to determine if gRNA expressed *in trans* could rescue the CA-SP1 processing defect of the LD4 mutant, and if there was any effect on processing and release of wild type virions. The experiment was also performed with constructs expressing the 5'UTR alone, with or without the packaging signal, to observe if this region is responsible for the effects observed when gRNA is expressed *in trans*. The requirement for ALIX was also tested by including a mutant lacking its YPX_nL binding motif.

5.3. Results

5.3.1. Overexpression of gRNA accelerates CA-SP1 processing

To produce a construct which can be transcribed to produce gRNA but without being translated, a single mutation was made in the Gag start codon, changing it from AUG to AAG. Following transfection and Western blotting with an anti-CA antibody a truncated version of Gag was observed (Figure 35 lane 2). This was due to the presence of an alternative start codon in Gag (Poon et al., 2002). Therefore, a premature stop codon was introduced into the CA domain of Gag. Western blotting confirmed that the additional mutation blocked Gag expression (Figure 35 lane 3), so this construct was used to express gRNA *in trans* for subsequent experiments.

Figure 35 – Expression profiles of wild type Gag and Gag mutants



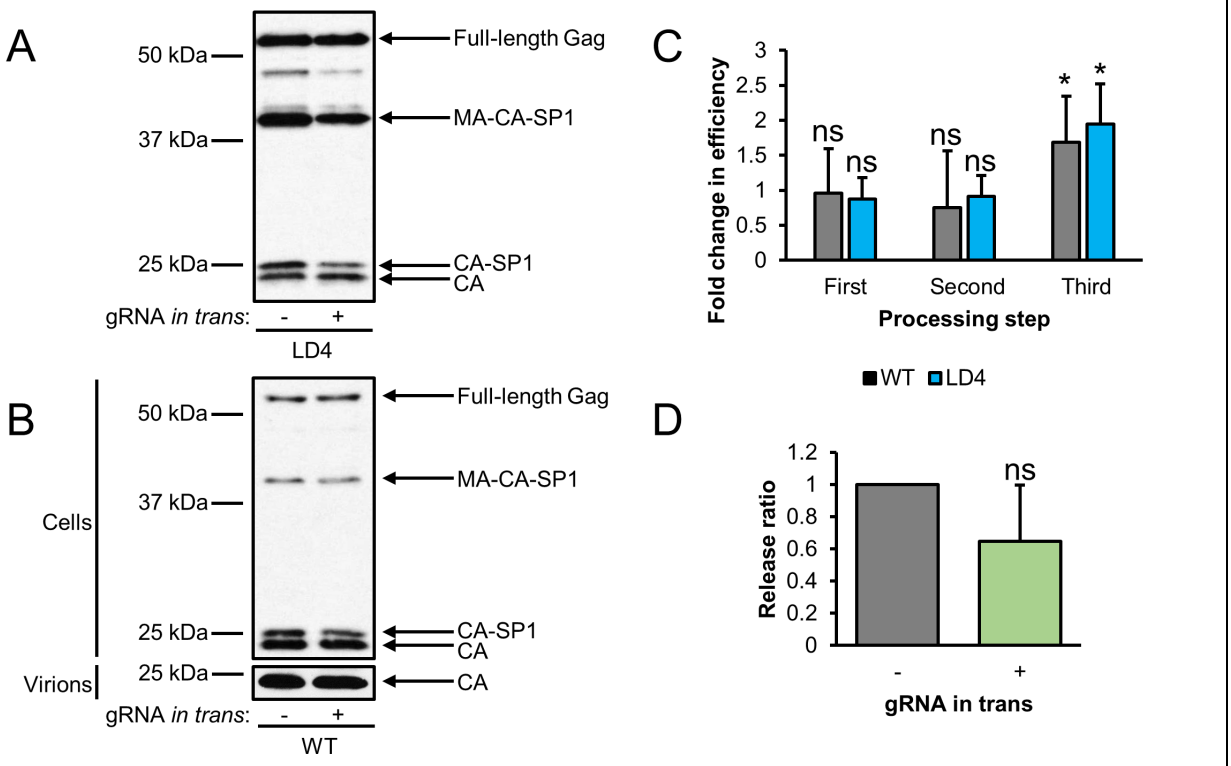
Western blot of cell lysates following transfection with the indicated proviruses. AAG = single mutant containing Gag start codon mutation; AAG/CA = double mutant containing Gag start codon mutation and frameshift in CA.

gRNA was expressed *in trans* to observe the effects on defective CA-SP1 processing caused by the LD4 mutation (Figure 36A) and on CA-SP1 processing in the wild type virus (Figure 36B).

Expression of wild type gRNA *in trans* with the LD4 mutant led to elevation of CA-SP1 processing (Figure 36**A**), but more intriguingly, CA-SP1 processing in the wild type virus was also accelerated (Figure 36B). Quantitation revealed that in both the wild type and LD4 mutant the effect of gRNA expression *in trans* on processing was specific to the final step of cleavage (Figure 36**C**).

However, no difference was observed in the amount of virus present in the supernatant for the wild type virus (Figure 36**D**). Virion samples for the LD4 mutant were unavailable, so it wasn't possible to assess the effect of gRNA on release of the LD4 mutant.

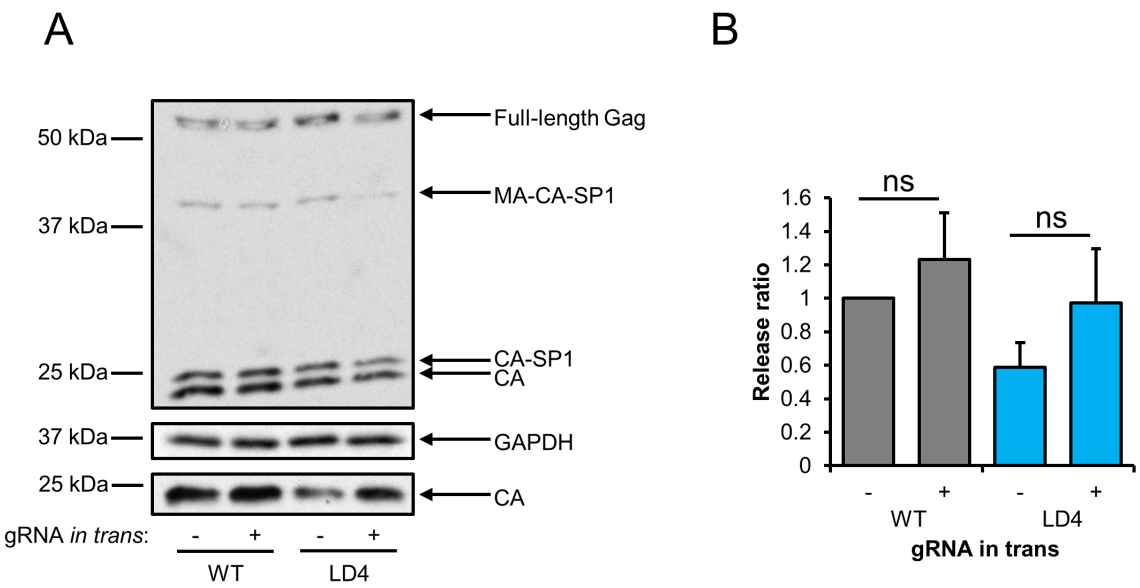
Figure 36 – Effects of in *trans* gRNA expression on Gag processing and virus release – first experiment



Western blot of cell lysates (A, B) following transfection with the indicated proviruses with or without overexpression of gRNA. C) Quantitation of the fold change in the efficiency of each step of processing upon gRNA overexpression, performed as described in Figure 18. The graph shows the mean values from at least 3 independent experiments. Error bars represent standard deviation. D) Quantitation of wild type virus release ratios without and with gRNA overexpression. The graph shows the mean values from 4 independent experiments. Error bars represent standard deviation.

The experiment was therefore repeated to determine the effect of overexpression on both WT and LD4 virus release, and to compare the virus constructs side by side in the same gels (Figure 37).

Figure 37 – Effects of *in trans* gRNA expression on Gag processing and virus release – second experiment



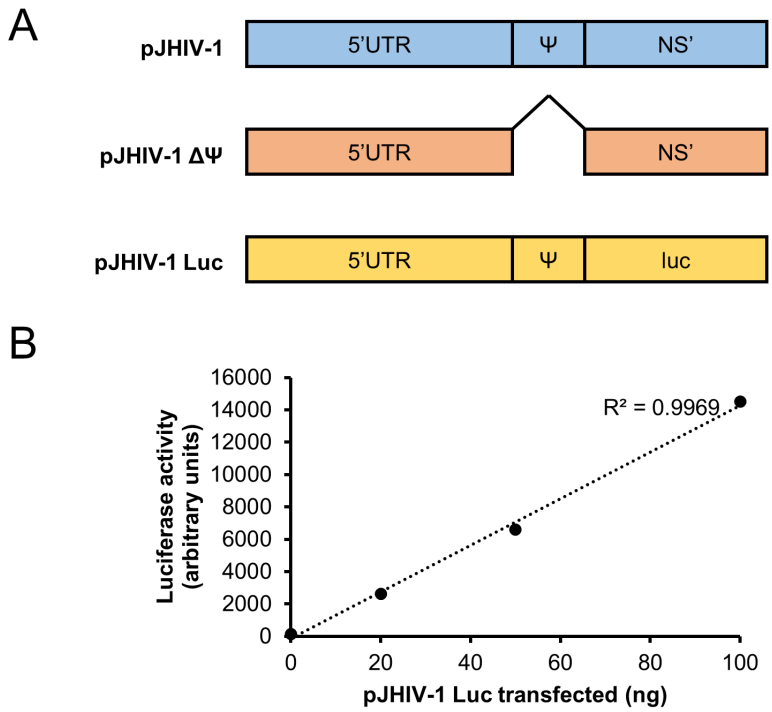
A) Western blot of cell lysates following transfection with the indicated proviruses, with or without overexpression of gRNA. B) Quantitation of mean release ratios relative to WT, from three independent experiments. Error bars represent standard deviation.

Upon repeating the experiment, the effect of gRNA on CA-SP1 → CA processing was far less pronounced, although in Figure 37A above some degree of acceleration is observed for the LD4 mutant. In agreement with the first experiment, the effect of overexpressing gRNA on WT virus release is minimal. However, LD4 mutant release rises from approximately 60% relative to WT, to be on par with WT release. This result isn't statistically significant ($p = 0.13$) but with further repeats a trend may become more apparent ($n = 3$).

5.3.2. Expression of wild type 5'UTR containing the packaging signal in trans is sufficient to accelerate CA-SP1 processing

It has been shown that mutating SL1 is sufficient to delay CA-SP1 processing (Figure 15). An interesting question then arose as to whether *in trans* expression of the 5'UTR alone would be able to reverse the Gag cleavage defect caused by the LD4 mutation. To investigate this, constructs expressing the full length 5'UTR (pJHIV-1, pJHIV-1 Luc) or truncated 5'UTR lacking stem-loops 1-4 (pJHIV-1 ΔΨ) were expressed *in trans.* pJHIV-1 and pJHIV-1 ΔΨ express the 5'UTR upstream of a truncated influenza NS gene, and pJHIV-1 Luc expresses the 5'UTR upstream of luciferase (Figure 38A). First, to confirm expression, a range of transfections were performed with increasing amounts of pJHIV-1 Luc, and luciferase expression was measured. The signal increased linearly with transfection amount showing that production of truncated viral RNA correlates with the amount of construct transfected (Figure 38B).

Figure 38 – 5'UTR constructs and expression of luciferase from pJHIV-1 Luc

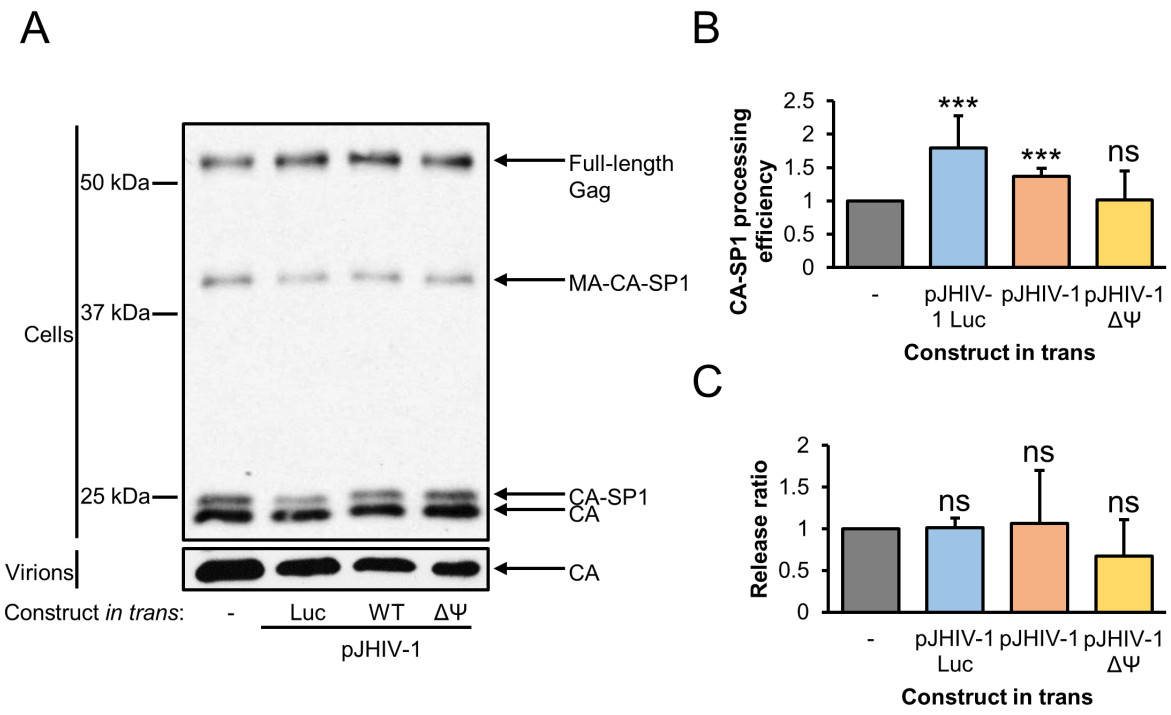


A) Cartoon representation of the composition of the 5'UTR constructs. B) Graph showing the effect of transfecting increasing amounts of pJHIV-1 Luc construct on the expression of luciferase.

Having confirmed 5'UTR expression the constructs were expressed alongside the wild type provirus, and the effect on Gag processing was assessed by Western blotting (Figure 39).

Expression of pJHIV-1 Luc or pJHIV-1 *in trans* was sufficient to accelerate CA-SP1 processing to a similar extent to that seen with the gRNA constructs (Figure 39B). However, when stem-loops 1-4 were deleted the rate of CA-SP1 processing remained unchanged. There was no significant difference in the amount of virus released under any of the conditions (Figure 39C).

Figure 39 – Effects of in trans 5'UTR expression on Gag processing and virus release

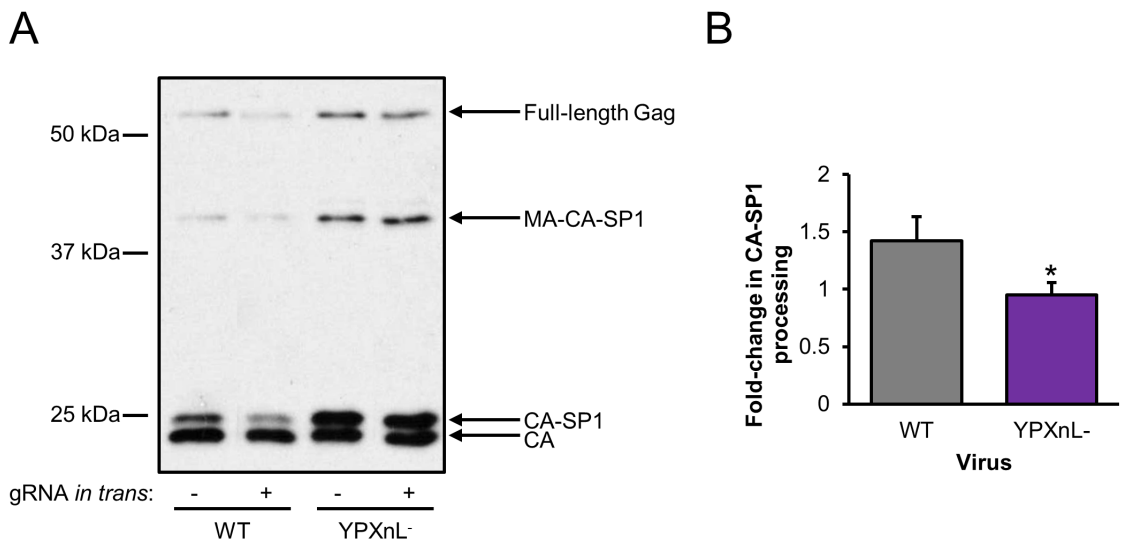


A) Western blot of cell lysates and purified supernatants following transfection with the indicated proviruses. B) Quantitation of relative CA-SP1 processing efficiency without and with 5'UTR overexpression. The graph shows the mean values from at least 4 independent experiments. Error bars represent standard deviation C) Quantitation of relative release ratios without and with 5'UTR overexpression. The graph shows the mean values from at least 3 independent experiments. Error bars represent standard deviation.

5.3.3. Acceleration of CA-SP1 processing by gRNA overexpression requires YPXnL motif

The acceleration of processing caused by *in trans* expression of gRNA expressing-constructs containing the packaging signal is reminiscent of the effect observed upon overexpression of ALIX (Figure 31). To test if this acceleration is ALIX-dependent, the experiment was repeated with the YPX_nL^- mutant (Figure 40).

Figure 40 – Effect of in trans gRNA expression on Gag processing in a YPX_nL^- mutant



A) Western blot of cell lysates following transfection with the indicated proviruses with or without gRNA overexpression. B) Quantitation of the fold-change in CA-SP1 processing upon overexpression of gRNA. The graph shows the mean values from 3 independent experiments. Error bars represent standard deviation.

Intriguingly, the acceleration of processing observed by expressing gRNA *in trans* Gag was blocked when the YPX_nL motif was mutated.

5.4. Discussion

The experiments described in this chapter were performed to further clarify the role of gRNA in virus maturation and release. It was hypothesised that the role of gRNA is to assist in the recruitment of ESCRT proteins to ensure the preferential release of gRNA-containing virions. If this is the case, then increasing the concentration of gRNA in the cell should influence CA-SP1 processing and virus release.

Consistent with this, the defect in processing caused by the LD4 mutation could be rescued by expression of wild type gRNA *in trans* (Figure 36A), and virus release also appeared to be rescued (Figure 37). These observations may be due to the wild type gRNA out-competing the LD4 mutant gRNA for packaging, resulting in the formation of 'hybrid' virions which are essentially wild type in composition and phenotype.

Intriguingly, expression of gRNA *in trans* with a wild type provirus also accelerated processing (Figure 36), reminiscent of the acceleration observed by overexpression of ALIX (Figure 31) and introduction of compensatory mutations into Gag (Figure 18). The packaging signal is responsible for this effect since expression of the 5'UTR alone is sufficient, providing stem-loops 1-4 are intact (Figure 39). This region of the 5'UTR contains multiple Gag binding sites (Section 1.12) (Kenyon et al., 2015), so it

appears likely that binding of Gag to gRNA stimulates these phenotypes. Acceleration was also dependent on an intact YPX_nL binding motif (Figure 40).

Despite enhanced processing, there was no effect on wild type virus release (Figure 36B & C, Figure 37, and Figure 39A). After these experiments had been performed two highly relevant papers were published by Becker *et al* and Dilley *et al* (Becker and Sherer, 2017; Dilley et al., 2017). They investigated the effects of altering gRNA subcellular localisation (Becker and Sherer, 2017) and abundance (Dilley et al., 2017) on formation and release of Gag complexes. In both studies experiments were performed where non-translatable gRNA was provided *in trans*. Becker *et al* co-expressed codon-optimised Gag-GFP, derived from a Gag expressing construct lacking PR, and an MS2-labelled gRNA construct containing a Gag start codon mutation. There was no significant difference in virus release when gRNA was expressed *in trans*, but the subcellular localisation of Gag changed. Dilley *et al* expressed a gRNA construct containing a frameshift in Gag and a mixture of CeFP-labelled and unlabelled helper construct. The helper construct expressed all genes except for *Env* and contained a mutated packaging signal to prevent *in cis* encapsidation. When gRNA, or a mini construct containing just the 5'UTR, was expressed *in trans*, there was a 2.5-fold increase in virus release. This effect was dependent on the presence of the packaging signal and the effect size was inversely proportional to the concentration of Gag, suggesting that the role of gRNA in assembly and release is more important at lower, physiologically relevant, Gag concentrations.

The authors hypothesised that gRNA acts as a nucleation point for particle assembly. If this is correct, the possibility that some of the enhanced CA-SP1 processing observed in these experiments was due to enhanced nucleation of assembly cannot be excluded, although no significant increases in virus release were observed (Figure 36 and Figure 39), in agreement with Becker *et al* (Becker and Sherer, 2017). However, the finding that the YPX_nL motif was required for this phenomenon (Figure 40) combined with data from the previous chapter showing that acceleration of CA-SP1 processing by compensatory mutations was late domain-dependent (Figure 32 and Figure 33) suggests that gRNA binding to Gag *per se* is not sufficient to accelerate processing, but that ALIX and potentially other ESCRT factors are involved.

The reason for the discrepancy in virus release is unclear, but the experimental conditions varied a great deal between those experiments and the ones performed here. Dilley *et al* noted that the effect of gRNA expression on release was lessened at higher concentrations of Gag, so differences in Gag concentration, due to different transfection protocols or promoter activities, may be one explanation. In this study and the Becker *et al* study 1:1 and 3:1 ratios of gRNA:Gag were used, whereas the ratio used by Dilley *et al* is not stated. Perhaps a much higher concentration of gRNA is required to see a significant effect on virus release.

When the experiment in Figure 36 was repeated, the CA-SP1 → CA cleavage acceleration effect was far less apparent. The reason for this is unclear, but it is difficult to imagine that the accelerated processing upon gRNA overexpression observed in Figure 36 is artefactual, as the trend is clear in that blot and others, such as the blot in Figure 39, produced at the time.

Noting the above caveat, these results suggest that gRNA regulates CA-SP1 processing in a psi- and ALIX- dependent manner, supporting the hypothesis that the NC domain of Gag recruits ALIX in a psi-dependent manner to promote ESCRT-III recruitment to the budding neck.

6. Conclusions

6.1. Overview

The experiments described in this thesis were designed to further explore the link between genome dimerisation and Gag processing in HIV-1, with a focus on the effect of mutations in the packaging signal region of the 5'UTR. Previous work by others has shown that delaying or inhibiting Gag processing hinders maturation of the gRNA dimer (Darlix et al., 1990; Muriaux et al., 1996), however less is known about the reverse process – the effect (if any) of genome dimerisation in influencing the rate of Gag processing. Work in HIV-2 showed that mutation of a critical GGAG motif hindered dimerisation and also delayed a step of Gag processing, resulting in the formation of virions with abnormal morphology (L'Hernault et al., 2007), and in HIV-1 it was shown that gross mutation of SL1 also affects both processes (Liang et al., 1999a), but that work was performed using substantial deletions of SL1 that likely have a major effect on the tertiary structure of the region surrounding SL1.

6.2. Investigating the link between the packaging signal and regulation of Gag processing (Chapters 3 and 4)

Using more targeted mutation of SL1, I found an example of a mutation (Stem B) that disturbed genome dimerisation whilst leaving the rate of the various steps of Gag processing unaffected. This suggests that efficient gRNA dimerisation is not required for efficient Gag processing, in contrast with the findings in HIV-2. Gag processing was studied by Western blotting at fixed time points post-transfection. Results were reproducible, but an effort was made to complement this with pulse-chase metabolic labelling to study processing over time. This was challenging due to a high level of non-specific background caused by inefficient immunoprecipitation, but the results seem to support those obtained using Western blots (see Appendix C).

The previous work in HIV-2 and HIV-1 highlighted the ability of replicating viruses to compensate for disabling packaging signal mutations, enabling restoration of Gag processing and (in the case of HIV-2) dimerisation. Using molecular cloning I introduced compensatory mutations previously shown to rescue the Gag processing defects of large SL1 mutants (Liang et al., 1999a) into the SL1 mutants tested in this study. Not only did these compensatory mutations correct defective CA-SP1 processing the Loop B and LD4 mutants, but they also accelerated processing in the case of the wild type, DIS and Stem B mutants. The acceleration was specific to the final stage of Gag processing, when CA-SP1 is cleaved into CA and SP1, an event that triggers the dramatic structural change in virions whereby the immature spherical core becomes a mature conical one (Wiegers et al., 1998). Taken together these novel findings suggest that particular components of SL1 play an important role in coordinating Gag processing. Currently much more is known about the effect of SL1 on genome dimerisation, but these results shed some light on its effects on Gag processing.

CA-SP1 processing defects are a hallmark of 'late domain' mutants ('late' in this context referring to the late stage of the virus replication cycle that the mutations disrupt). The late domain (also known as p6) is responsible for recruitment of ESCRT components ALIX and TSG101, enabling the virus to hijack this host process to enable it to bud from the cell. It was therefore of interest to determine if the mutations in SL1 affect virus release and the recruitment of ESCRT components. This was found to be the case, with the previously studied LD4 mutation causing an approximately four-fold reduction in the amount of virus released into the supernatant. To exclude the possibility that this was caused by excessive degradation of virions rather than reduced release, a stability assay was conducted. Although no statistically significant difference in stability was observed, there was an indication that stability may play a role, although it is unlikely that this is the sole reason for the reduced virus release.

In attempting to formulate a hypothesis to explain the link between mutations in the packaging signal and the hijacking of host ESCRT machinery, I came across papers that showed the ESCRT protein ALIX binds to NC (the region of Gag that binds the packaging signal) (Popov et al., 2008), and that it does so in an RNA-dependent manner (Sette et al., 2012). A rescue experiment was performed in which ALIX was overexpressed in the context of the $PTAP^-$ mutant control, which rescues release, and a mutant combining this mutation with the LD4 SL1 mutation. The results showed that $PTAP^-$/LD4 was rescued with 50% the efficiency of $PTAP^-$. Interestingly, the LD4 mutation acted similarly to late domain and NC mutations in that it also blocked the acceleration of CA-SP1 caused by ALIX overexpression.

To establish a link between these observations and the compensatory mutation findings discussed in Chapter 3, the compensatory mutations were tested for their ability to rescue CA-SP1 processing in the context of a disabled ESCRT system (using both late domain mutagenesis and siRNA knockdown of ALIX and TSG101). These experiments revealed a requirement for TSG01 and ALIX and their interaction with Gag for the compensatory mutations to act in this way.

Results from Chapter 3 and Chapter 4 provide the first evidence, to my knowledge, that the packaging signal of HIV-1 interacts with the host ESCRT pathway, and that disturbing the packaging signal through mutagenesis affects recruitment of ALIX, potentially causing the aberrant Gag processing previously observed by others and confirmed here but hereto unexplained. These results do not confirm that the disruption of ESCRT recruitment is responsible for the disturbed Gag processing, but they show a correlations, and the reductions in CA-SP1 processing efficiency and virus release are likely to be due to altered ESCRT binding, as these defects are hallmarks of interference with ESCRT recruitment (Fisher et al., 2007; Garrus et al., 2001; Huang et al., 1995; Meng et al., 2015; Morita et al., 2011)

6.3. Proposed model explaining the link between the packaging signal and effects on Gag processing and virus release

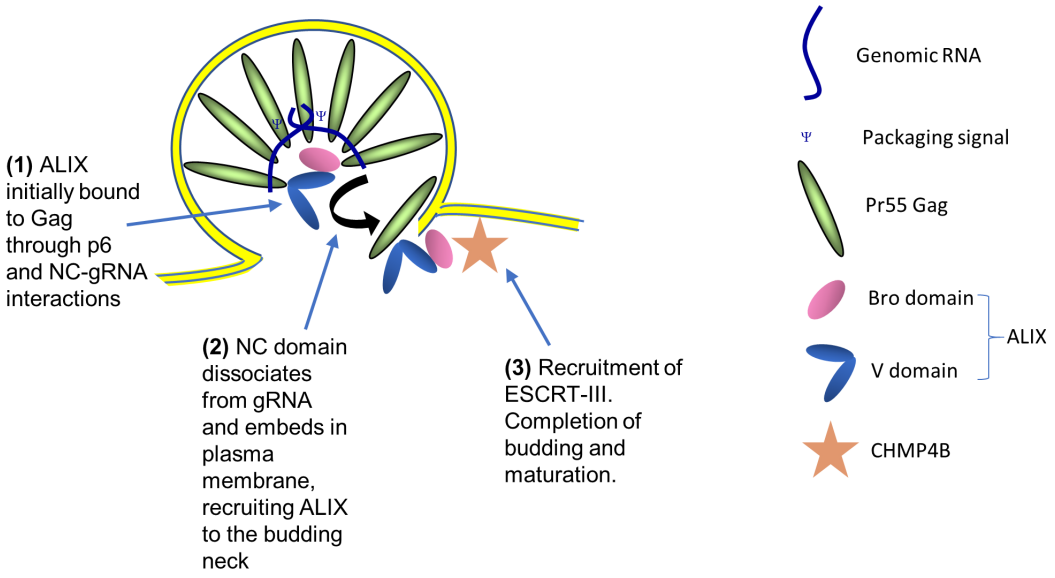
The results discussed above suggests that the function of ALIX is directly affected by mutation of SL1. The model from Sette *et al* for the role of ALIX in virus assembly proposes that a subset of Gag molecules switch binding preference from an NC-RNA interaction to an NC-lipid one to enable co-anchoring with the Bro1 domain of ALIX via a lipid intermediate in the budding neck (Figure 24).

The finding that RNA can act as a bridge between the two molecules *in vitro* (Sette et al., 2012) is not accounted for in this model. It is also not known what would determine which subset of the 2400 Gag molecules (Carlson et al., 2008) in the virion make the switch from RNA binding to membrane binding.

If RNA acts as a bridge to recruit the Bro1 domain of ALIX to NC *in vivo*, it could be gRNA or any spliced viral or cellular RNA as NC possesses specific and non-specific RNA binding capability. However, the finding that SL1 mutation results in a phenotype consistent with reduced ESCRT recruitment in Chapter 3, and that ALIX function is disrupted by SL1 mutation in Chapter 4 suggests that gRNA is involved in ALIX recruitment during viral assembly.

Each virion contains only two copies of gRNA, compared to 2400 Gag molecules. This stoichiometry means that only a few NC domains would be bound to the packaging signal region in the 5'UTR. An attractive hypothesis would be that ALIX is recruited to the subset of NC domains that are bound to this region. gRNA binding to NC may maintain the zinc fingers in a configuration preferentially recognised by Bro1. The RNA bridge would then be traded for a lipid one causing ALIX to localise to the budding neck. ALIX could then recruit ESCRT-III to sever the budding neck and trigger virion maturation (Figure 41).

Figure 41 – Model for role of gRNA in virus maturation and release



ALIX initially anchors onto the p6 domain of Gag in the immature virion. Packaging of gRNA results in the formation of a specific contact between the NC domain of Gag and structures such SL1 in the 5'UTR. This conformation is recognised by the Bro1 domain of ALIX enabling a second contact to be made. Migration of the Gag molecule to the budding neck results in replacement of the gRNA bridge for a lipid one. ESCRT-III proteins are then recruited by ALIX, triggering virion maturation and release.

This model accounts for the observations made in the experiments performed in this study and others. Mutation of SL1 or NC reduces the formation of Bro1-NC complexes mediated by gRNA (and lipids in the case of NC mutation). This reduces migration of ALIX to the budding neck and subsequent ESCRT-III recruitment required for virus maturation and release. Overexpression of ALIX increases the formation of gRNA-mediated Bro1-NC complexes, reversing these phenotypes.

An additional role for gRNA in the less well understood NC-TSG101 interaction cannot be excluded. It is possible that the LD4 mutation interferes with both the ALIX and TSG101 interactions with NC, as both ALIX and TSG101 recruit ESCRT-III (Garrus et al., 2001; Kim et al., 2005) and have been shown to bind to NC *in vitro* in the presence of RNA (Chamontin et al., 2015; Sette et al., 2012).

6.4. Investigating the effect of varying cellular gRNA levels (Chapter 5)

The experiments described in Chapter 5 were motivated by a desire to better understand the apparent requirement for coordination to occur between gRNA packaging and virus release. Work in the past decade has highlighted the degree of success with which HIV-1 packages its genome – over 90% of virions in one study contained gRNA (Chen et al., 2009). One hypothesis arising from this observation is that successful gRNA packaging is a checkpoint in the virus release process, ensuring preferential release of virions that have successfully packaged gRNA and are thus capable of replicating. This would

prevent the release of VLPs derived from multimerisation of Gag independently of gRNA, which would only serve to attract the attention of the immune system and would not result in productive infection.

Based on the results in Chapters 3 and 4 it seemed plausible that regulation of release could work by ESCRT recruitment (in particular ALIX) being gRNA-dependent. The first experiments in Chapter 5 assessed the effect of overexpressing gRNA or a construct containing the 5'UTR only, on Gag processing and virus release.

Initial experiments suggested that overexpression of gRNA specifically accelerates the CA-SP1 step of Gag processing (analogous to the effect of the compensatory mutations investigated in Chapter 4), in both wild type and an LD4 mutant virus. To clarify the effect on virus release the experiment was repeated. Although the result was not statistically significant, there was an indication that overexpressing gRNA may contribute to restoring the defect in virus release caused by the LD4 mutation.

The next set of experiments looked at the effect of overexpressing the 5'UTR in the context of a non-viral RNA, to confirm that this is the region responsible for the effects on processing and release. This indeed turned out to be the case, with both constructs containing a complete packaging signal accelerating processing, and the one lacking the packaging signal failing to do so. In agreement with the gRNA overexpression results, WT virus release was not changed.

Considering the proposed model presented in Figure 41, gRNA overexpression may be increasing the opportunity for interaction between NC and ALIX, accelerating ESCRT recruitment and therefore Gag processing and virus release. For the final experiment performed in Chapter 5, gRNA was overexpressed in the wild type virus as before, and in the YPXnL^{-} mutant, which lacks the primary ALIX binding site. Intriguingly, this mutation appeared to prevent the gRNA-mediated acceleration of processing, further adding weight to the argument that acceleration of processing depends on the interplay between ALIX and gRNA.

6.5. Future experiments

The results presented here show correlations, and suggest causations, between various events in the late stages of HIV-1 assembly, but more conclusive evidence remains elusive. A number of different experimental techniques were attempted during the project, but these were either unsuccessful or not reproducible. In others' hands these may help to make the findings here more robust.

Perhaps the most obvious next step would be to demonstrate that the physical interaction between Gag and ALIX is dependent on SL1. I attempted to do this using electrophoretic mobility shift assay experiments, but these were hampered by a lack of sensitivity. The rationale was to incubate purified Gag and purified ALIX *in vitro* in the presence of either WT SL1 or SL1 containing mutations such as the

Loop B and LD4 mutations described in this report. I would then separate these samples using electrophoresis and observe the effect, if any, of the mutation in SL1 on the mobility of Gag and ALIX. Whilst I was able to observe specific bands, the low intensity made it difficult to determine if the mutation was resulting in a different phenotype. These experiments were attempted using non-radiolabelled probes, so using a radiolabelled assay should help to improve sensitivity. Encouragingly, work by other members of the lab has produced some positive preliminary results, which suggest that mutation of SL1 does indeed hinder the interaction between Gag and ALIX.

Another approach would be to use immunofluorescence or fluorescence microscopy to study the interaction in cells. Again, preliminary work was conducted using immunofluorescence, and whilst Gag was readily observed in cells, ALIX was difficult to label so it was decided not to pursue this line of work. If set up correctly, this system could be used to look at co-localisation of Gag and ALIX in cells expressing either WT or SL1-mutant virus, and also the effect of any differences in co-localisation on the timing and extent of virus budding. By using fluorescently-labelled proteins, fluorescence microscopy could be performed, although this introduces the risk of the fusion protein having different biological function to its natural counterpart.

A less direct but equally interesting study could look at the recruitment of ESCRT proteins into wild type and SL1-mutant virions, This could be done using mass spectrometry, for example. It would be of interest to measure the levels of ALIX, TSG101 and other ESCRT proteins incorporated into WT and SL1-mutant virions.

Given the parallels between the effects of the compensatory mutations and MI resistance mutations, it would be interesting to test whether the compensatory mutations provide any resistance against MIs and vice versa.

A study by Ogawa et al (Ogawa et al., 2015) found a novel interaction between the SP1 protein (also known as p2; produced by the key Gag processing step discussed in this thesis – the proteolytic cleavage of CA-SP1 into CA and SP1) and mitochondrial protein important for ATP synthesis. They found reduced intracellular ATP production in MT-4 cells and MDMs in cells infected with HIV-1 lacking functional SP1, and that this in turn had negative effect on reverse transcription. I performed some preliminary experiments attempting to measure the levels of ATP and the effect on reverse transcription in cells following infection with WT and SL1 mutant viruses, but the experiments were difficult to reproduce.

6.6. Importance of this work

The hypotheses discussed in this thesis represent a new way of looking at the role played by the 5'UTR in the late stages of HIV-1 particle assembly and release, and the results, although not always conclusive, tend to support these hypotheses and help to frame future research questions. While much

has been known about the effect of Gag processing on HIV-1 dimer formation and maturation, little is known about the reverse side of the coin – the effect of dimerisation on Gag processing and subsequent virus release. As we continue to grapple for a HIV cure, increased understanding of its molecular mechanisms provides us with new drug targets. RNA structures such as SL1 are amongst the most highly conserved regions of the virus genome, which, in a virus with a great capacity to evolve resistance, makes them desirable drug targets (Le Grice, 2015). Preclinical research into the activity of RNA aptamers directed against the 5'UTR highlights the potential of this approach (Sánchez-Luque et al., 2014). By showing that SL1 is involved in processes other than packaging, the desirability of targeting this region of the genome is increased, as disruption of multiple key processes in virus assembly reduces the risk of the virus evolving resistance and increases the potential potency of such drugs.

7. Appendices

Appendix A: Buffer and solution recipes

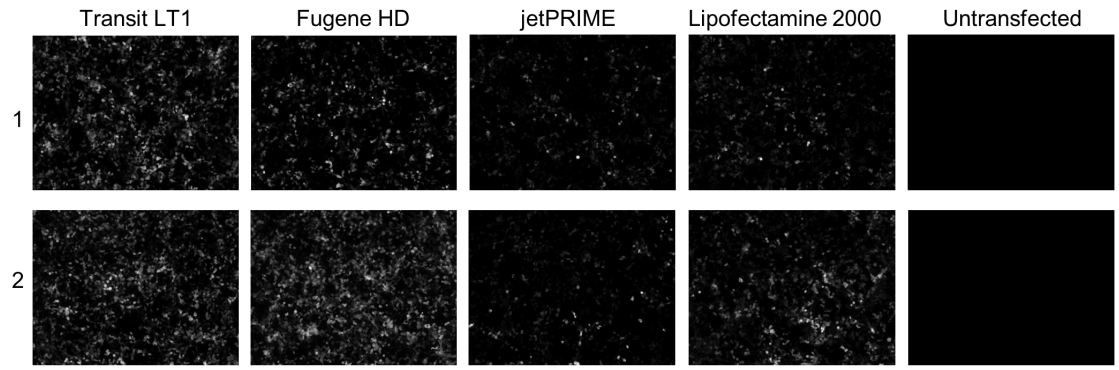
Table 10 – buffers and solutions used

Solution	Recipe
dNTP mix	10mM dATP, 10mM dCTP, 10mM dGTP, 10mM dTTP (all Promega)
ELISA coating antibody solution	10 µg/ml anti-p24 antibody (Aalto Bio Reagents), 0.1M $NaHCO_3$
ELISA secondary antibody solution	2% (w/v) milk powder, 79.5% (v/v) 1X TBS, 20% (v/v) sheep serum, 0.5% 1XTBS-10% Tween, 1/16000 (v/v) alkaline phosphatase-conjugated anti-p24 antibody (Aalto Bio Reagents)
Gel drying solution	30% (v/v) methanol, 5% (v/v) glycerol
Gel fixing solution	40% (v/v) methanol, 10% (v/v) acetic acid
High stringency wash buffer	15 mM NaCl, 1.5 mM trisodium citrate, 0.1% (w/v) SDS
Laemmli buffer (2X)	20% (v/v) glycerol, 0.1% (w/v) bromophenol blue, 100 mM 2-mercaptoethanol, 2% (w/v) SDS, 125 mM Tris-HCl pH 6.8
Low stringency wash buffer	300 mM NaCl, 30 mM trisodium citrate, 0.1% (w/v) SDS
LB agar	1% (w/v) tryptone, 1% (w/v) NaCl, 0.5% (w/v) yeast extract, 1.5% (w/v) agar
LB medium	1% (w/v) tryptone, 1% (w/v) NaCl, 0.5% (w/v) yeast extract
Proteinase K extraction buffer	50 mM Tris-HCl pH 7.5, 100 mM NaCl, 10 mM EDTA, 1% (w/v) SDS, 100 µg/ml yeast tRNA (Sigma-Aldrich) and 100 µg/ml proteinase K solution (Thermo Fisher)
Radioimmunoprecipitation assay (RIPA) buffer	1% (v/v) Triton X-100, 1% (w/v) sodium deoxycholate, 0.1% (w/v) SDS, 100 mM NaCl, 50 mM Tris-HCl pH 7.5
Saline-sodium citrate (SSC) buffer (20X)	3M NaCl, 300 mM trisodium citrate adjusted to pH 7.0 with HCl
Sodium dodecyl sulfate-polyacrylamide gel electrophoreses (SDS-PAGE) gel	12-15% (v/v) ProtoGel acrylamide mix (National Diagnostics), 375 mM Tris (pH 8.8), 0.1 % (w/v) SDS, 0.1 % ammonium persulfate, 0.01 % TEMED
Tris-borate-EDTA (TBE) buffer	89 mM tris, 89 mM boric acid, 2 mM EDTA
Tris-buffered saline (TBS)	50 mM Tris-HCl pH 7.5, 150 mM NaCl
Western blot running buffer	25 mM tris, 250 mM glycine, 0.1% (w/v) SDS
Western blot transfer buffer	25 mM tris, 192 mM glycine, 20% (v/v) ethanol

Appendix B: Comparison of transfection reagents

Four transfection reagents were tested for their transfection efficiency in 293T cells by transfecting a plasmid expressing eGFP into 293T cells (Figure 42).

Figure 42 – Comparison of transfection reagent efficiency

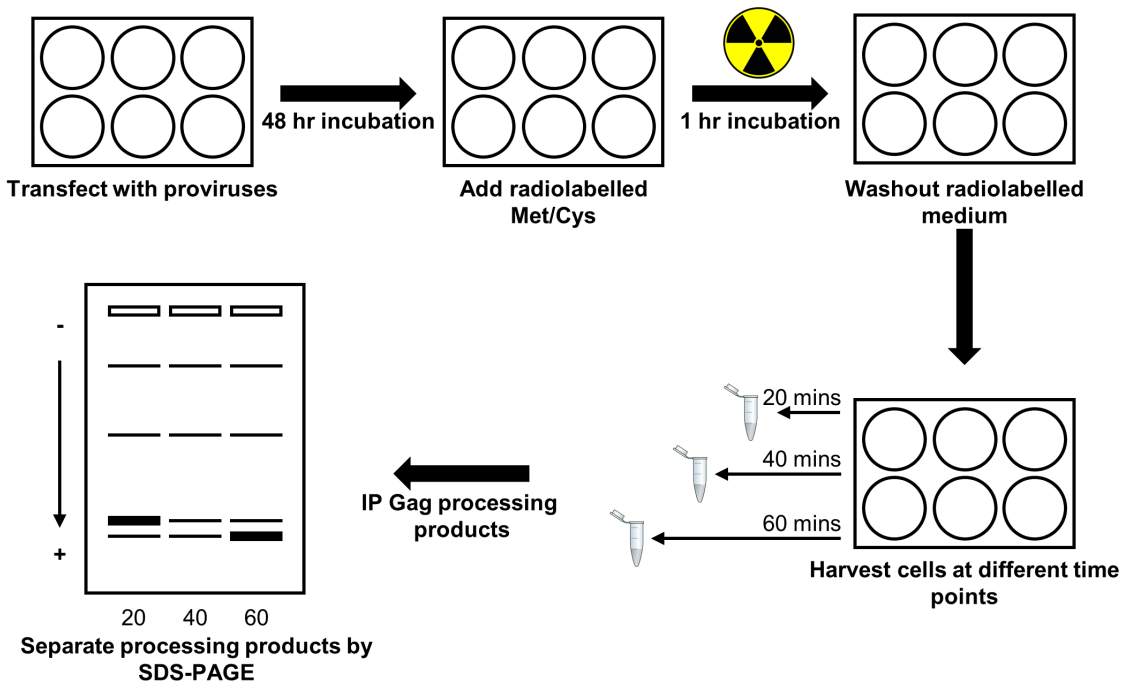


Transit-LT1, Fugene HD (Promega), jetPRIME (Polyplus-transfection) and Lipofectamine 2000 (Thermo Fisher) were compared by following the manufacturers' recommended protocols to transfect 200 ng plasmid DNA per well of an 80% confluent 24-well plate. Two independent transfections were performed. 48 hours post-transfection the supernatants were discarded, and the cells were analysed under a fluorescence microscope

*Trans*it-LT1 and Fugene HD were comparable in terms of their transfection efficiency, whereas jetPRIME and Lipofectamine 2000 performed poorly. Since *Trans*it-LT1 appeared to be marginally superior to Fugene HD this transfection reagent was used for transfections in the study.

Appendix C: Pulse-chase assessment of Gag processing

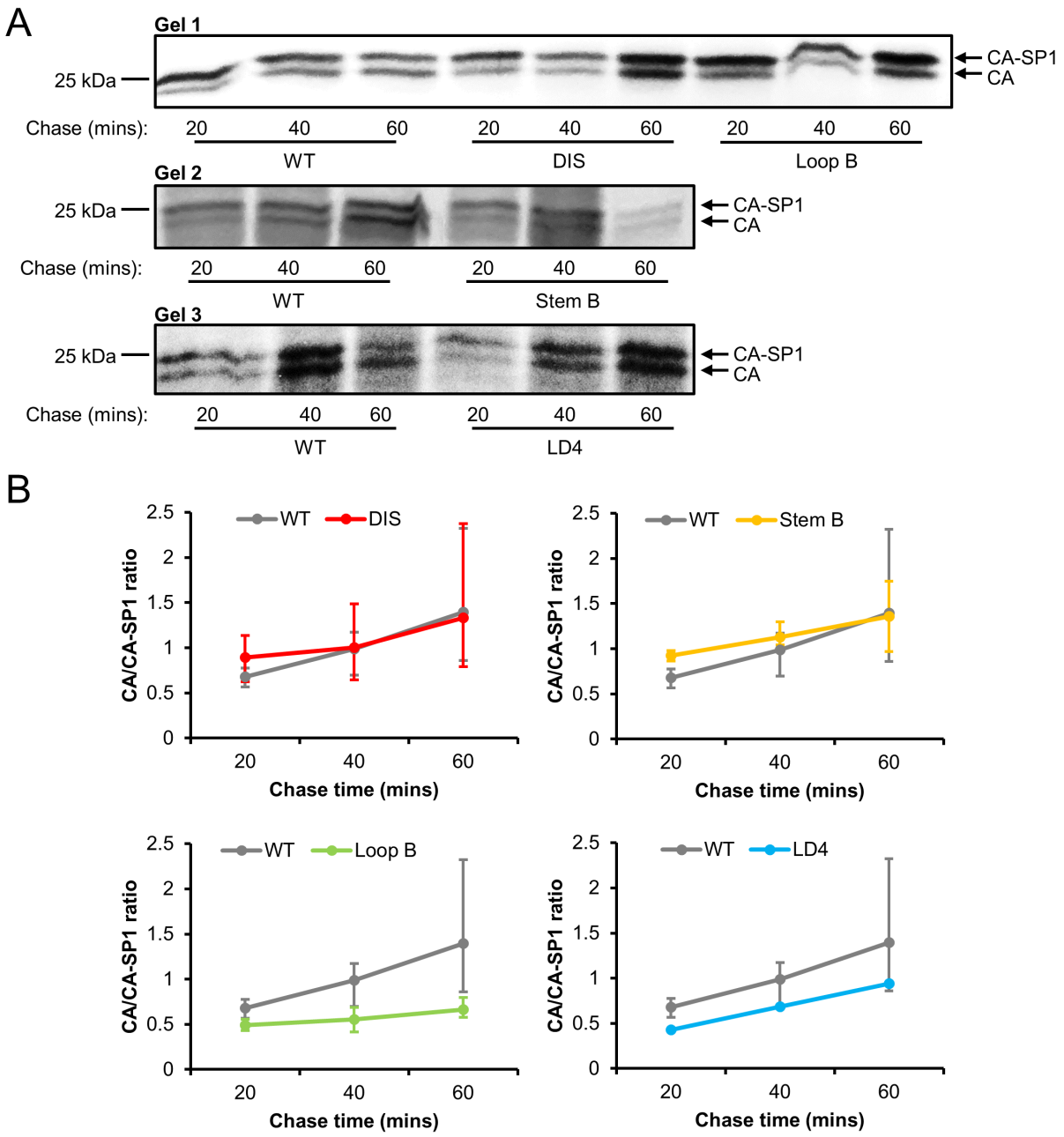
Figure 43 – Pulse-chase experimental setup



The medium on wild type and mutant provirus-transfected cells was replaced with medium containing radiolabelled methionine and cysteine for one hour. This was then removed, and cells were harvested at regular intervals, before immunoprecipitating Gag and its processing products. Processing products were separated by SDS-PAGE before exposure to film.

Due to high levels of background from other radiolabelled proteins it was not possible to perform quantitation of the full-length Gag and MA-CA-SP1 bands. Examples of gels showing the CA-SP1 and CA bands for each mutant following 20, 40 and 60-minute chases are shown in Figure 44A, with quantitation of the average CA/CA-SP1 ratios over time for each mutant in Figure 44B. The quality of the blots is poor due to faint bands so interpretation of the quantified plots must be performed with caution, however by inspecting the blots it appears that processing in the DIS mutant is no different to wild type, whilst processing in the Loop B and LD4 mutants is delayed especially at the 60-minute time point. The bands in the Stem B blot are extremely faint but appear to exhibit wild type-like processing.

Figure 44 – Pulse-chase analysis of CA-SP1 processing in cells



A) Representative gels showing the processing of CA-SP1 over time. It was not possible to run all mutants on the same gel so 3 independent gels using different mutants are shown. B) The graphs show the mean ratio of CA/CA-SP1 at each time point for each mutant, overlaid onto the wild type graph. At least 2 independent experiments were performed for the wild type, DIS, Stem B and Loop B viruses, and 1 experiment was performed for the LD4 mutant. Error bars represent range.

Appendix D: Phenotypic properties of SL1 mutants

Table 11 – phenotypes of SL1 mutants

Mutant name	Infectivity (% WT) (Cell type)	Dimerisation (% WT) (Cell type)	Packaging (% WT) (Cell type)	Reference
Large deletions				
Δ241-253	Delayed replication (MT-4)	-	< 20% (SW480)	(Kim et al., 1994)
Δ243-277	0.1% (SupT1)	-	100% (COS7) 20% (SupT1)	(Paillart et al., 1996)
Δ248-270	0.1% (SupT1)	-	100% (COS7) 20% (SupT1)	(Paillart et al., 1996)
ΔSL1	< 10% (HOS)	-	19% (293T)	(Clever and Parslow, 1997)
Δ248-261	< 1% (MT-4)	57% (COS7)	57% (COS7)	(Laughrea et al., 1997)
A1	Lethal (Jurkat-*tat*)	-	-	(Harrison et al., 1998)
A2	Severely impaired (Jurkat-*tat*)	-	5% (Jurkat-*tat*)	(Harrison et al., 1998)
A3	Moderately impaired (Jurkat-*tat*)	-	50% (Jurkat-*tat*)	(Harrison et al., 1998)
LD3/Δ241-256	< 1% (MT-2)	53% (COS7)	20% (COS7) 30% (COS7)	(Liang et al., 1998; Shen et al., 2000)
LD4[1]	< 1% (MT-2)	-	70% (COS7)	(Liang et al., 1999b, 1998)
LD5	Extremely defective (MT-2)	-	-	(Liang et al., 1999a)
ΔDIS	Unable to replicate (MT-2, CBMC) Unable to establish persistent infection (Jurkat) 30% (Hela-CD4-LTR-β-Gal)	50% (COS7)	65% (COS7)	(Russell et al., 2003)
NLACGCGT	WT-like (PBMC) Unable to replicate (SupT1)	100% (but more diffuse dimers) (293T)	50% (293T)	(Hill et al., 2003)
NLGCGCGC	WT-like (PBMC) Unable to replicate (SupT1)	100% (but more diffuse dimers) (293T)	50% (293T)	(Hill et al., 2003)

ΔSL1	< 5% (PM-1)	56% (293T)	50% (293T)	(Ristic and Chin, 2010)
Palindrome mutants				
GC1	Delayed replication (SupT1)	100% (SupT1)	50% (SupT1)	(Berkhout, 1996)
GC5	Delayed replication (SupT1)	100% (SupT1)	50% (SupT1)	(Berkhout, 1996)
G-loop	< 10% (HOS)	-	83% (293T)	(Clever and Parslow, 1997)
GCGCUC	< 10% (HOS)	-	42% (293T)	(Clever and Parslow, 1997)
KL^{-}	< 25% (C8166)	Reduced but not abolished (C8166)	100% (C8166)	(Haddrick et al., 1996)
S257-259	1% (SupT1)	-	100% (COS7) 50% (SupT1)	(Paillart et al., 1996)
S255-263	10% (SupT1)	-	100% (COS7) 50% (SupT1)	(Paillart et al., 1996)
ACS^{-}	< 1% (MT-4)	not quantified but described as less dimeric (COS7)	24% (COS7)	(Laughrea et al., 1997)
GGCG^{2}	< 5% (MT-4)	59% (COS7)	60% (COS7) 78% (COS7)	(Laughrea et al., 1999; Shen et al., 2000)
GGCC	-	97% (COS7)	-	(Shen et al., 2000)
ΔLoop	-	50% (COS7)	90% (COS7)	(Russell et al., 2003)
ΔDIS	-	100% (293T)	-	(Song et al., 2007)
C258G^{2}	-	100% (293T)	-	(Song et al., 2007)
Stem mutants				
dS.1	< 10% (HOS)	-	11% (293T)	(Clever and Parslow, 1997)
mS.1	< 50% (HOS)	-	42% (293T)	(Clever and Parslow, 1997)
S1	-	-	80% (HeLa)	(McBride and Panganiban, 1996)
S1	-	-	39% (293)	(McBride and Panganiban, 1997)
A11	Severely impaired (Jurkat-*tat*)	-	20% (Jurkat-*tat*)	(Harrison et al., 1998)
A12	Moderately impaired (Jurkat-*tat*)	-	60% (Jurkat-*tat*)	(Harrison et al., 1998)

A13	100% (Jurkat-*tat*)	-	100% (Jurkat-*tat*)	(Harrison et al., 1998)
Δ243-247	< 0.1% (MT-4)	56% (COS7)	60% (COS7)	(Laughrea et al., 1999; Shen et al., 2000)
Δ248-256	< 0.1% (MT-4)	53% (COS7)	70% (COS7)	(Laughrea et al., 1999; Shen et al., 2000)
Δ243-246[3]	< 0.1% (COS-7)	59% (COS7)	25% (COS7)	(Shen et al., 2001)
Δ274-277	< 0.1% (COS-7)	59% (COS7)	40% (COS7)	(Shen et al., 2001)
Loop mutants				
ΔB.1	< 10% (HOS)	-	14% (293T)	(Clever and Parslow, 1997)
mB.1	< 10% (HOS)	-	18% (293T)	(Clever and Parslow, 1997)
Δ271-273[4]	< 1% (COS-7)	59% (COS7)	44% (COS7)	(Shen et al., 2001)
Δ247	< 5% (COS-7)	59% (COS7)	65% (COS7)	(Shen et al., 2001)

Numbers in superscript format indicate mutants that are equivalent to the LD4 (1), DIS (2), Stem B (3) and Loop B (4) mutants used in this study. Mutant names that include nucleotide positions use the numbering system shown in Figure 3 and not the system used in Materials and Methods.

8. Bibliography

Abbink, T.E.M., Berkhout, B., 2003. A novel long distance base-pairing interaction in human immunodeficiency virus type 1 RNA occludes the Gag start codon. J. Biol. Chem. 278, 11601–11. doi:10.1074/jbc.M210291200

Abd El-Wahab, E.W., Smyth, R.P., Mailler, E., Bernacchi, S., Vivet-Boudou, V., Hijnen, M., Jossinet, F., Mak, J., Paillart, J.-C., Marquet, R., 2014. Specific recognition of the HIV-1 genomic RNA by the Gag precursor. Nat. Commun. 5, 4304. doi:10.1038/ncomms5304

Accola, M.A., Strack, B., Gottlinger, H.G., 2000. Efficient Particle Production by Minimal Gag Constructs Which Retain the Carboxy-Terminal Domain of Human Immunodeficiency Virus Type 1 Capsid-p2 and a Late Assembly Domain. J. Virol. 74, 5395–5402. doi:10.1128/JVI.74.12.5395-5402.2000

Adamson, C.S., Ablan, S.D., Boeras, I., Goila-Gaur, R., Soheilian, F., Nagashima, K., Li, F., Salzwedel, K., Sakalian, M., Wild, C.T., Freed, E.O., 2006. In Vitro Resistance to the Human Immunodeficiency Virus Type 1 Maturation Inhibitor PA-457 (Bevirimat). J. Virol. 80, 10957–10971. doi:10.1128/JVI.01369-06

Alcamí, J., Laín de Lera, T., Folgueira, L., Pedraza, M.A., Jacqué, J.M., Bachelerie, F., Noriega, A.R., Hay, R.T., Harrich, D., Gaynor, R.B., 1995. Absolute dependence on kappa B responsive elements for initiation and Tat-mediated amplification of HIV transcription in blood CD4 T lymphocytes. EMBO J. 14, 1552–60.

Alfadhli, A., Still, A., Barklis, E., 2009. Analysis of human immunodeficiency virus type 1 matrix binding to membranes and nucleic acids. J. Virol. 83, 12196–12203. doi:10.1128/JVI.01197-09

Alizon, S., Magnus, C., 2012. Modelling the course of an HIV infection: Insights from ecology and evolution. Viruses 4, 1984–2013. doi:10.3390/v4101984

Anderson, E.C., Lever, A.M.L., 2006. Human immunodeficiency virus type 1 Gag polyprotein modulates its own translation. J. Virol. 80, 10478–86. doi:10.1128/JVI.02596-05

Archin, N.M., Margolis, D.M., 2014. Emerging strategies to deplete the HIV reservoir. Curr. Opin. Infect. Dis. 27, 29–35. doi:10.1097/QCO.0000000000000026

Arhel, N.J., Souquere-Besse, S., Munier, S., Souque, P., Guadagnini, S., Rutherford, S., Prévost, M.-C., Allen, T.D., Charneau, P., 2007. HIV-1 DNA Flap formation promotes uncoating of the pre-integration complex at the nuclear pore. EMBO J. 26, 3025–37. doi:10.1038/sj.emboj.7601740

Arrigo, S.J., Weitsman, S., Zack, J.A., Chen, I.S., 1990. Characterization and expression of novel singly spliced RNA species of human immunodeficiency virus type 1. J Virol 64, 4585–4588.

Babst, M., Katzmann, D.J., Estepa-Sabal, E.J., Meerloo, T., Emr, S.D., 2002a. ESCRT-III: An Endosome-Associated Heterooligomeric Protein Complex Required for MVB Sorting. Dev. Cell 3, 271–282. doi:10.1016/S1534-5807(02)00220-4

Babst, M., Katzmann, D.J., Snyder, W.B., Wendland, B., Emr, S.D., 2002b. Endosome-associated complex, ESCRT-II, recruits transport machinery for protein sorting at the multivesicular body. Dev. Cell 3, 283–289. doi:10.1016/S1534-5807(02)00219-8

Barouch, D.H., 2009. Challenges in the Development of an HIV-1 Vaccine. Nature 455, 613–619. doi:10.1038/nature07352.Challenges

Barre-Sinoussi, F., Chermann, J.-C., Rey, F., Nugeyre, M.T., Chamaret, S., Gruest, J., Dauguet, C., Axler-Blin, C., Zinet-Brun, F., Rouzioux, C., Rozenbaum, W., Montagnier, L., 1983. Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). Science (80-.). 220, 868–871.

Bayro, M.J., Ganser-Pornillos, B.K., Zadrozny, K.K., Yeager, M., Tycko, R., 2016. Helical Conformation in the CA-SP1 Junction of the Immature HIV-1 Lattice Determined from Solid-State NMR of Virus-like Particles. J. Am. Chem. Soc. jacs.6b07259. doi:10.1021/jacs.6b07259

Becker, J.T., Sherer, N.M., 2017. Subcellular Localization of HIV-1 gag-pol mRNAs Regulates Sites of Virion Assembly. J. Virol. 91, 1–22.

Bell, N.M., Lever, A.M.L., 2013. HIV Gag polyprotein: processing and early viral particle assembly. Trends Microbiol. 21, 136–44. doi:10.1016/j.tim.2012.11.006

Bendjennat, M., Saffarian, S., 2016. The Race against Protease Activation Defines the Role of ESCRTs in HIV Budding. PLOS Pathog. 12, e1005657. doi:10.1371/journal.ppat.1005657

Berkhout, B., 1996. Role of the DIS Hairpin in Replication of Human Immunodeficiency Virus Type 1. J. Virol. 70, 6723–6732.

Berkhout, B., van Wamel, J.L., 2000. The leader of the HIV-1 RNA genome forms a compactly folded tertiary structure. RNA 6, 282–95. doi:10.1017/S1355838200991684

Bleck, M., Itano, M.S., Johnson, D.S., Thomas, V.K., North, A.J., Bieniasz, P.D., Simon, S.M., 2014. Temporal and spatial organization of ESCRT protein recruitment during HIV-1 budding. Proc. Natl. Acad. Sci. U. S. A. 111, 12211–6. doi:10.1073/pnas.1321655111

Braaten, D., Franke, E.K., Luban, J., 1996. Cyclophilin A is required for an early step in the life cycle of human immunodeficiency virus type 1 before the initiation of reverse transcription. J. Virol. 70, 3551–3560.

Brass, A.L., Dykxhoorn, D.M., Benita, Y., Yan, N., Engelman, A., Xavier, R.J., Lieberman, J., Elledge, S.J., 2008. Identification of Host Proteins Required for HIV Infection Through a Functional Genomic Screen. Science (80-.). 319, 921–926. doi:10.1126/science.1152725

Briggs, J.A.G., Wilk, T., Welker, R., Kräusslich, H.-G., Fuller, S.D., 2003. Structural organization of authentic, mature HIV-1 virions and cores. EMBO J. 22, 1707–1715. doi:10.1093/emboj/cdg143

Burke, D.S., 1997. Recombination in HIV: An Important Viral Evolutionary Strategy. Emerg. Infect. Dis. 3, 253–259. doi:10.3201/eid0303.970301

Campbell, S., Rein, A., 1999. In vitro assembly properties of human immunodeficiency virus type 1 Gag protein lacking the p6 domain. J. Virol. 73, 2270–2279.

Campbell, S., Vogt, V.M., 1995. Self-assembly in vitro of purified CA-NC proteins from Rous sarcoma virus and human immunodeficiency virus type 1. J. Virol. 69, 6487–6497.

Campsteijn, C., Vietri, M., Stenmark, H., 2016. Novel ESCRT functions in cell biology: spiraling out of control? Curr. Opin. Cell Biol. 41, 1–8. doi:10.1016/j.ceb.2016.03.008

Carlson, L.-A., Briggs, J.A.G., Glass, B., Riches, J.D., Simon, M.N., Johnson, M.C., Müller, B., Grünewald, K., Kräusslich, H.-G., 2008. Three-dimensional analysis of budding sites and released virus suggests a revised model for HIV-1 morphogenesis. Cell Host Microbe 4, 592–9. doi:10.1016/j.chom.2008.10.013

Carlson, L.A., Hurley, J.H., 2012. In Vitro Reconstitution of the Ordered Assembly of the ESCRT Machinery at Membrane-Bound HIV-1 Gag Clusters. Proc Natl Acad Sci U S A 109, 16928–16933. doi:10.1073/pnas.1211759109/-/DCSupplemental.www.pnas.org/cgi/doi/10.1073/pnas.1211759109

Carlton, J.G., Martin-Serrano, J., 2007. Parallels between cytokinesis and retroviral budding: a role for the ESCRT machinery. Science 316, 1908–12. doi:10.1126/science.1143422

Chamontin, C., Rassam, P., Ferrer, M., Racine, P.J., Neyret, A., Laine, S., Milhiet, P.E., Mougel, M., 2015. HIV-1 nucleocapsid and ESCRT-component Tsg101 interplay prevents HIV from turning into a DNA-containing virus. Nucleic Acids Res. 43, 336–347. doi:10.1093/nar/gku1232

Chen, J., Grunwald, D., Sardo, L., Galli, A., Plisov, S., Nikolaitchik, O.A., Chen, D., Lockett, S., Larson, D.R., Pathak, V.K., Hu, W.-S., 2014. Cytoplasmic HIV-1 RNA is mainly transported by diffusion in the presence or absence of Gag protein. Proc. Natl. Acad. Sci. U. S. A. 111, E5205-13. doi:10.1073/pnas.1413169111

Chen, J., Nikolaitchik, O.A., Singh, J., Wright, A., Bencsics, C.E., Coffin, J.M., Ni, N., Lockett, S., Pathak, V.K., Hu, W.-S., 2009. High efficiency of HIV-1 genomic RNA packaging and heterozygote

formation revealed by single virion analysis. Proc. Natl. Acad. Sci. U. S. A. 106, 13535–40. doi:10.1073/pnas.0906822106

Chen, J., Rahman, S.A., Nikolaitchik, O.A., Grunwald, D., Sardo, L., Burdick, R.C., Plisov, S., Liang, E., Tai, S., Pathak, V.K., Hu, W.-S., 2015. HIV-1 RNA genome dimerizes on the plasma membrane in the presence of Gag protein. Proc. Natl. Acad. Sci. U. S. A. 113, E201-8. doi:10.1073/pnas.1518572113

Chen, N.-Y., Zhou, L., Gane, P.J., Opp, S., Ball, N.J., Nicastro, G., Zufferey, M., Buffone, C., Luban, J., Selwood, D., Diaz-Griffero, F., Taylor, I., Fassati, A., 2016. HIV-1 capsid is involved in post-nuclear entry steps. Retrovirology 13, 28. doi:10.1186/s12977-016-0262-0

Cherepanov, P., Maertens, G., Proost, P., Devreese, B., Van Beeumen, J., Engelborghs, Y., De Clercq, E., Debyser, Z., 2003. HIV-1 integrase forms stable tetramers and associates with LEDGF/p75 protein in human cells. J. Biol. Chem. 278, 372–381. doi:10.1074/jbc.M209278200

Cheung, M.C., 2005. AIDS-Related Malignancies: Emerging Challenges in the Era of Highly Active Antiretroviral Therapy. Oncologist 10, 412–426. doi:10.1634/theoncologist.10-6-412

Chin, C.R., Perreira, J.M., Savidis, G., Portmann, J.M., Aker, A.M., Feeley, E.M., Smith, M.C., Brass, A.L., 2015. Direct Visualization of HIV-1 Replication Intermediates Shows that Capsid and CPSF6 Modulate HIV-1 Intra-nuclear Invasion and Integration. Cell Rep. 13, 1717–1731. doi:10.1016/j.celrep.2015.10.036

Chin, M.P.S., Rhodes, T.D., Chen, J., Fu, W., Hu, W.-S., 2005. Identification of a major restriction in HIV-1 intersubtype recombination. Proc. Natl. Acad. Sci. U. S. A. 102, 9002–7. doi:10.1073/pnas.0502522102

Chojnacki, J., Staudt, T., Glass, B., Bingen, P., Engelhardt, J., Anders, M., Schneider, J., Müller, B., Hell, S.W., Kräusslich, H.G., 2012. Maturation-dependent HIV-1 surface protein redistribution revealed by fluorescence nanoscopy. Science (80-.). 338, 524–528. doi:10.1126/science.1226359

Chojnacki, J., Waithe, D., Carravilla, P., Huarte, N., Galiani, S., Enderlein, J., Eggeling, C., 2017. Envelope glycoprotein mobility on HIV-1 particles depends on the virus maturation state. Nat. Commun. 8. doi:10.1038/s41467-017-00515-6

Chukkapalli, V., Oh, S.J., Ono, A., 2010. Opposing mechanisms involving RNA and lipids regulate HIV-1 Gag membrane binding through the highly basic region of the matrix domain. Proc. Natl. Acad. Sci. U. S. A. 107, 1600–1605. doi:10.1073/pnas.0908661107

Chukkapalli, V., Ono, A., 2011. Molecular determinants that regulate plasma membrane association

of HIV-1 Gag. J. Mol. Biol. 410, 512–524. doi:10.1016/j.jmb.2011.04.015

Cimarelli, A., Sandin, S., Höglund, S., Luban, J., 2000. Basic residues in HIV-1 nucleocapsid promote virion assembly via interaction with RNA. J. Virol. 74, 3046–3057.

Ciuffi, A., Llano, M., Poeschla, E., Hoffmann, C., Leipzig, J., Shinn, P., Ecker, J.R., Bushman, F., 2005. A role for LEDGF/p75 in targeting HIV DNA integration. Nat. Med. 11, 1287–1289. doi:10.1038/nm1329

Clever, J.L., Eckstein, D.A., Parslow, T.G., 1999. Genetic dissociation of the encapsidation and reverse transcription functions in the 5' R region of human immunodeficiency virus type 1. J. Virol. 73, 101–9.

Clever, J.L., Parslow, T.G., 1997. Mutant human immunodeficiency virus type 1 genomes with defects in RNA dimerization or encapsidation. J. Virol. 71, 3407–14.

Coffin, J.M., 1979. Structure, Replication, and Recombination of Retrovirus Genomes: Some Unifying Hypotheses. J. Gen. Virol. 42, 1–26.

Cosnefroy, O., Murray, P.J., Bishop, K.N., 2016. HIV-1 capsid uncoating initiates after the first strand transfer of reverse transcription. Retrovirology 13. doi:10.1186/s12977-016-0292-7

Craigie, R., Bushman, F.D., 2012. HIV DNA integration. Cold Spring Harb. Perspect. Med. 2. doi:10.1101/cshperspect.a006890

Cullen, B.R., 2003. Nuclear mRNA export: Insights from virology. Trends Biochem. Sci. 28, 419–424. doi:10.1016/S0968-0004(03)00142-7

Cullen, B.R., 2000. Connections between the processing and nuclear export of mRNA: evidence for an export license? Proc. Natl. Acad. Sci. U. S. A. 97, 4–6. doi:10.1073/PNAS.97.1.4

D'arc, M., Ayouba, A., Esteban, A., Learn, G.H., Boué, V., Liegeois, F., Etienne, L., Tagg, N., Leendertz, F.H., Boesch, C., Madinda, N.F., Robbins, M.M., Gray, M., Cournil, A., Ooms, M., Letko, M., Simon, V.A., Sharp, P.M., Hahn, B.H., Delaporte, E., Mpoudi Ngole, E., Peeters, M., 2015. Origin of the HIV-1 group O epidemic in western lowland gorillas. Proc. Natl. Acad. Sci. 112, E1343–E1352. doi:10.1073/pnas.1502022112

Dalgleish, A.G., Beverley, P.C.L., Clapham, P.R., Crawford, D.H., Greaves, M.F., Weiss, R.A., 1984. The CD4 (T4) antigen is an essential component of the receptor for the AIDS retrovirus. Nature 312, 763–767. doi:10.1038/312763a0

Damgaard, C.K., Dyhr-Mikkelsen, H., Kjems, J., 1998. Mapping the RNA binding sites for human immunodeficiency virus type-1 Gag and NC proteins within the complete HIV-1 and -2

untranslated leader regions. Nucleic Acids Res. 26, 3667–3676. doi:10.1093/nar/26.16.3667

Darlix, J.-L., Gabus, C., Nugeyre, M.-T., Clavel, F., Barré-Sinoussi, F., 1990. Cis elements and Trans-acting factors involved in the RNA dimerization of the human immunodeficiency virus HIV-1. J. Mol. Biol. 216, 689–699. doi:10.1016/0022-2836(90)90392-Y

Das, A.T., Klaver, B., Klasens, B.I., van Wamel, J.L., Berkhout, B., 1997. A conserved hairpin motif in the R-U5 region of the human immunodeficiency virus type 1 RNA genome is essential for replication. J. Virol. 71, 2346–56.

De Iaco, A., Santoni, F., Vannier, A., Guipponi, M., Antonarakis, S., Luban, J., 2013. TNPO3 protects HIV-1 replication from CPSF6-mediated capsid stabilization in the host cell cytoplasm. Retrovirology 10, 20. doi:10.1186/1742-4690-10-20

Dean, M., Carrington, M., Winkler, C., Huttley, G.A., Smith, M.W., Allikmets, R., Goedert, J.J., Buchbinder, S.P., Vittinghoff, E., Gomperts, E., Donfield, S., Vlahov, D., Kaslow, R., Saah, A., Rinaldo, C., Detels, R., O'Brien, S.J., 1996. Genetic restriction of HIV-1 infection and progression to AIDS by a deletion allele of the CKR5 structural gene. Science (80-.). 273, 1856–1862. doi:10.1126/science.273.5283.1856

Demirov, D.G., Orenstein, J.M., Freed, E.O., 2002. The Late Domain of Human Immunodeficiency Virus Type 1 p6 Promotes Virus Release in a Cell Type-Dependent Manner. J. Virol. 76, 105–117. doi:10.1128/JVI.76.1.105

Di Nunzio, F., Danckaert, A., Fricke, T., Perez, P., Fernandez, J., Perret, E., Roux, P., Shorte, S., Charneau, P., Diaz-Griffero, F., Arhel, N.J., 2012. Human Nucleoporins Promote HIV-1 Docking at the Nuclear Pore, Nuclear Import and Integration. PLoS One 7. doi:10.1371/journal.pone.0046037

Dilley, K.A., Nikolaitchik, O.A., Galli, A., Burdick, R.C., Levine, L., Li, K., Rein, A., Pathak, V.K., Hu, W.-S., 2017. Interactions Between HIV-1 Gag and Viral RNA Genome Enhance Virion Assembly. J. Virol. JVI.02319-16. doi:10.1128/JVI.02319-16

Dingwall, C., Ernberg, I., Gait, M.J., Green, S.M., Heaphy, S., Karn, J., Lowe, A.D., Singh, M., Skinner, M.A., Valerio, R., 1989. Human immunodeficiency virus 1 tat protein binds trans-activation-responsive region (TAR) RNA in vitro. Proc. Natl. Acad. Sci. U. S. A. 86, 6925–9. doi:10.1073/pnas.86.18.6925

Dorfman, T., Luban, J., Goff, S.P., Haseltine, W. a, Göttlinger, H.G., 1993. Mapping of functionally important residues of a cysteine-histidine box in the human immunodeficiency virus type 1 nucleocapsid protein. J. Virol. 67, 6159–6169.

Dussupt, V., Javid, M.P., Abou-Jaoudé, G., Jadwin, J.A., De La Cruz, J., Nagashima, K., Bouamr, F., 2009. The nucleocapsid region of HIV-1 gag cooperates with the PTAP and LYPX nL late domains to recruit the cellular machinery necessary for viral budding. PLoS Pathog. 5. doi:10.1371/journal.ppat.1000339

Dussupt, V., Sette, P., Bello, N.F., Javid, M.P., Nagashima, K., Bouamr, F., 2011. Basic residues in the nucleocapsid domain of Gag are critical for late events of HIV-1 budding. J. Virol. 85, 2304–15. doi:10.1128/JVI.01562-10

Eckwahl, M.J., Telesnitsky, A., Wolin, S.L., 2016. Host RNA packaging by retroviruses: A newly synthesized story. MBio 7, 1–8. doi:10.1128/mBio.02025-15

Faria, N.R., Rambaut, A., Suchard, M.A., Baele, G., Bedford, T., Ward, M.J., Tatem, A.J., Sousa, J.D., Arinaminpathy, N., Pepin, J., Posada, D., Peeters, M., Pybus, O.G., Lemey, P., 2014. The early spread and epidemic ignition of HIV-1 in human populations. Science (80-.). 346, 56–61. doi:10.1126/science.1256739

Fenner, 2016. Fenner's Veterinary Virology, 5th ed.

Ferrer, M., Clerté, C., Chamontin, C., Basyuk, E., Lainé, S., Hottin, J., Bertrand, E., Margeat, E., Mougel, M., 2016. Imaging HIV-1 RNA dimerization in cells by multicolor super-resolution and fluctuation microscopies. Nucleic Acids Res. gkw511. doi:10.1093/nar/gkw511

Finkelshtein, D., Werman, A., Novick, D., Barak, S., Rubinstein, M., 2013. LDL receptor and its family members serve as the cellular receptors for vesicular stomatitis virus. Proc. Natl. Acad. Sci. 110, 7306–7311. doi:10.1073/pnas.1214441110

Fisher, A.G., Collalti, E., Ratner, L., Gagllo, R.C., Wong-Staal, F., 1985. A molecular clone of HTLV-III with biological activity. Nature 316, 262–265. doi:10.1038/315279a0

Fisher, R.D., Chung, H.Y., Zhai, Q., Robinson, H., Sundquist, W.I., Hill, C.P., 2007. Structural and Biochemical Studies of ALIX/AIP1 and Its Role in Retrovirus Budding. Cell 128, 841–852. doi:10.1016/j.cell.2007.01.035

Freed, E.O., 2015. HIV-1 assembly, release and maturation. Nat. Rev. Microbiol. 13, 484–496. doi:10.1007/978-1-4614-7729-7

Fricke, T., Valle-Casuso, J.C., White, T.E., Brandariz-Nuñez, A., Bosche, W.J., Reszka, N., Gorelick, R., Diaz-Griffero, F., 2013. The ability of TNPO3-depleted cells to inhibit HIV-1 infection requires CPSF6. Retrovirology 10, 46. doi:10.1186/1742-4690-10-46

Fujii, K., Hurley, J.H., Freed, E.O., 2007. Beyond Tsg101: the role of Alix in "ESCRTing" HIV-1. Nat. Rev. Microbiol. 5, 912–916. doi:10.1038/nrmicro1790

Fujii, K., Munshi, U.M., Ablan, S.D., Demirov, D.G., Soheilian, F., Nagashima, K., Stephen, A.G., Fisher, R.J., Freed, E.O., 2009. Functional role of Alix in HIV-1 replication. Virology 391, 284–292. doi:10.1016/j.virol.2009.06.016

Gallo, R.C., Salahuddin, S.Z., Popovic, M., Shearer, G.M., Kaplan, M., Haynes, B.F., Palker, T.J., Redfield, R., Oleske, J., Safai, B., 1984. Frequent Detection and Isolation of Cytopathic Retroviruses. Science (80-.). 224, 500–503.

Garrus, J.E., Von Schwedler, U.K., Pornillos, O.W., Morham, S.G., Zavitz, K.H., Wang, H.E., Wettstein, D.A., Stray, K.M., Côté, M., Rich, R.L., Myszka, D.G., Sundquist, W.I., 2001. Tsg101 and the vacuolar protein sorting pathway are essential for HIV-1 budding. Cell 107, 55–65. doi:10.1016/S0092-8674(01)00506-2

Gheysen, D., Jacobs, E., de Foresta, F., Thiriart, C., Francotte, M., Thines, D., De Wilde, M., 1989. Assembly and release of HIV-1 precursor Pr55gag virus-like particles from recombinant baculovirus-infected insect cells. Cell 59, 103–112. doi:10.1016/0092-8674(89)90873-8

Göttlinger, H.G., Dorfman, T., Sodroski, J.G., Haseltine, W.A., 1991. Effect of mutations affecting the p6 gag protein on human immunodeficiency virus particle release. Proc. Natl. Acad. Sci. U. S. A. 88, 3195–3199.

Göttlinger, H.G., Sodroski, J.G., Haseltine, W.A., 1989. Role of capsid precursor processing and myristoylation in morphogenesis and infectivity of human immunodeficiency virus type 1. Proc. Natl. Acad. Sci. U. S. A. 86, 5781–5785. doi:10.1073/pnas.86.15.5781

Greatorex, J., Gallego, J., Varani, G., Lever, A.M.L., 2002. Structure and stability of wild-type and mutant RNA internal loops sfrom the SL-1 domain of the HIV-1 packaging signal. J. Mol. Biol. 322, 543–557. doi:10.1016/S0022-2836(02)00776-3

Grigorov, B., Décimo, D., Smagulova, F., Péchoux, C., Mougel, M., Muriaux, D., Darlix, J.-L., 2007. Intracellular HIV-1 Gag localization is impaired by mutations in the nucleocapsid zinc fingers. Retrovirology 4, 54. doi:10.1186/1742-4690-4-54

Guyader, M., Emerman, M., Sonigo, P., Clavel, F., Montagnier, L., Alizon, M., 1987. Genome organization and transactivation of the human immunodeficiency virus type 2. Nature 326, 662–669.

Haddrick, M., Lear, a L., Cann, a J., Heaphy, S., 1996. Evidence that a kissing loop structure facilitates genomic RNA dimerisation in HIV-1. J. Mol. Biol. 259, 58–68. doi:10.1006/jmbi.1996.0301

Han, Z., Madara, J.J., Liu, Y., Liu, W., Ruthel, G., Freedman, B.D., Harty, R.N., 2015. ALIX Rescues Budding of a Double PTAP/PPEY L-Domain Deletion Mutant of Ebola VP40: A Role for ALIX in

Ebola Virus Egress. J. Infect. Dis. 212, S138–S145. doi:10.1093/infdis/jiu838

Hanson, P.I., Roth, R., Lin, Y., Heuser, J.E., 2008. Plasma membrane deformation by circular arrays of ESCRT-III protein filaments. J. Cell Biol. 180, 389–402. doi:10.1083/jcb.200707031

Harrison, G.P., Miele, G., Hunter, E., Lever, A.M.L., 1998. Functional analysis of the core human immunodeficiency virus type 1 packaging signal in a permissive cell line. J. Virol. 72, 5886–96.

Hendrix, J., Baumgärtel, V., Schrimpf, W., Ivanchenko, S., Digman, M.A., Gratton, E., Kräusslich, H.-G., Müller, B., Lamb, D.C., 2015. Live-cell observation of cytosolic HIV-1 assembly onset reveals RNA-interacting Gag oligomers. J. Cell Biol. 210, 629–646. doi:10.1083/jcb.201504006

Henn, A., Flateau, C., Gallien, S., 2017. Primary HIV Infection: Clinical Presentation, Testing, and Treatment. Curr. Infect. Dis. Rep. 19. doi:10.1007/s11908-017-0588-3

Herrmann, C.H., Rice, a P., 1995. Lentivirus Tat proteins specifically associate with a cellular protein kinase, TAK, that hyperphosphorylates the carboxyl-terminal domain of the large subunit of RNA polymerase II: candidate for a Tat cofactor. J. Virol. 69, 1612–1620.

Hilditch, L., Towers, G.J., 2014. A model for cofactor use during HIV-1 reverse transcription and nuclear entry. Curr. Opin. Virol. 4, 32–36. doi:10.1016/j.coviro.2013.11.003

Hill, M.K., Shehu, M., Campbell, S.M., Poumbourios, P., Crowe, S.M., Mak, J., 2003. The Dimer Initiation Sequence Stem-Loop of Human Immunodeficiency Virus Type 1 Is Dispensable for Viral Replication in Peripheral Blood Mononuclear Cells. Microbiology 77, 8329–8335. doi:10.1128/JVI.77.15.8329

Hirsch, V.M., Olmsted, R.A., Murphey-Corb, M., Purcell, R.H., Johnson, P.R., 1989. An African primate lentivirus (SIVsm) closely related to HIV-2. Nature 339, 389–392. doi:10.1038/340301a0

Höglund, S., Öhagen, Å., Goncalves, J., Panganiban, A.T., Gabuzda, D., 1997. Ultrastructure of HIV-1 Genomic RNA. Virology 279, 271–279.

Holmes, C.B., Losina, E., Walensky, R.P., Yazdanpanah, Y., Freedberg, K., 2003. Review of human immunodeficiency virus type 1-related opportunistic infections in. Clin Infect Dis 36, 652–662. doi:10.1086/367655

Holmes, M., Zhang, F., Bieniasz, P.D., 2015. Single-Cell and Single-Cycle Analysis of HIV-1 Replication. PLoS Pathog. 11, 1–23. doi:10.1371/journal.ppat.1004961

Houzet, L., Paillart, J.-C., Smagulova, F., Maurel, S., Morichaud, Z., Marquet, R., Mougel, M., 2007. HIV controls the selective packaging of genomic, spliced viral and cellular RNAs into virions through different mechanisms. Nucleic Acids Res. 35, 2695–2704. doi:10.1093/nar/gkm153

Hu, W.-S., Hughes, S.H., 2012. HIV-1 reverse transcription. Cold Spring Harb. Perspect. Med. 2. doi:10.1101/cshperspect.a006882

Hu, W.-S., Temin, H.M., 1990. Genetic consequences of packaging two RNA genomes in one retroviral particle: pseudodiploidy and high rate of genetic recombination. Proc. Natl. Acad. Sci. U. S. A. 87, 1556–60. doi:10.1073/pnas.87.4.1556

Huang, M., Orenstein, J.M., Martin, M.A., Freed, E.O., 1995. p6Gag is required for particle production from full-length human immunodeficiency virus type 1 molecular clones expressing protease. J. Virol. 69, 6810–8.

Huet, T., Cheynier, R., Meyerhans, A., Roelants, G., Wain-Hobson, S., 1990. Genetic organization of a chimpanzee lentivirus related to HIV-1. Nature 345, 356–359.

Hulme, A.E., Perez, O., Hope, T.J., 2011. Complementary assays reveal a relationship between HIV-1 uncoating and reverse transcription. Proc. Natl. Acad. Sci. 108, 9975–9980. doi:10.1073/pnas.1014522108

Huthoff, H., Berkhout, B., 2001. Two alternating structures of the HIV-1 leader RNA. RNA 7, 143–57. doi:10.1017/S1355838201001881

Hutter, G., Nowak, D., Mossner, M., Ganepola, S., Müßig, A., Allers, K., Schneider, T., Hofmann, J., Kücherer, C., Blau, O., Blau, I.W., Hofmann, W.K., Thiel, E., 2009. Long-Term Control of HIV by CCR5 Delta32/ Delta32 Stem-Cell Transplantation. N. Engl. J. Med. 360, 692–698.

Isel, C., Karn, J., 1999. Direct evidence that HIV-1 Tat stimulates RNA polymerase II carboxyl-terminal domain hyperphosphorylation during transcriptional elongation. J. Mol. Biol. 290, 929–941. doi:10.1006/jmbi.1999.2933

Jacks, T., Power, M.D., Masiarz, F.R., Luciw, P.A., Barr, P.J., Varmus, H.E., 1988. Characterisation of ribosomal frameshifting in HIV-1 gag-pol expression. Nature 331, 280–283. doi:10.1038/332141a0

Jiang, M., Mak, J., Ladha, A., Cohen, E., Klein, M., Rovinski, B., Kleiman, L., 1993. Identification of tRNAs Incorporated into Wild-Type and Mutant Human Immunodeficiency Virus Type 1. J. Virol. 67, 3246–3253.

Jones, C.P., Cantara, W.A., Olson, E.D., Musier-Forsyth, K., 2014. Small-angle X-ray scattering-derived structure of the HIV-1 5’ UTR reveals 3D tRNA mimicry. Proc. Natl. Acad. Sci. U. S. A. 111, 3395–400. doi:10.1073/pnas.1319658111

Jouvenet, N., Bieniasz, P.D., Simon, S.M., 2008. Imaging the biogenesis of individual HIV-1 virions in live cells. Nature 454, 236–240. doi:10.1038/nature06998

Jouvenet, N., Simon, S.M., Bieniasz, P.D., 2009. Imaging the interaction of HIV-1 genomes and Gag during assembly of individual viral particles. Proc. Natl. Acad. Sci. U. S. A. 106, 19114–9. doi:10.1073/pnas.0907364106

Jouvenet, N., Zhadina, M., Bieniasz, P.D., Simon, S.M., 2011. Dynamics of ESCRT protein recruitment during retroviral assembly. Nat. Cell Biol. 13, 394–401. doi:10.1038/ncb2207

Kao, S.-Y., Calman, A.F., Luciw, P.A., Peterlin, B.M., 1987. Anti-termination of transcription within the long terminal repeat of HIV-1 by tat gene product. Nature 330, 489–493. doi:10.1038/330489a0

Kaplan, A.H., Manchester, M., Swanstrom, R., 1994. The Activity of the Protease of Human Immunodeficiency Virus Type 1 Is Initiated at the Membrane of Infected Cells before the Release of Viral Proteins and Is Required for Release To Occur with Maximum Efficiency. J. Virol. 68, 6782–6786.

Karacostas, V., Wolffe, E.J., Nagashima, K., Gonda, M.A., Moss, B., 1993. Overexpression of the HIV-1 Gag-Pol Polyprotein Results in Intracellular Activation of HIV-1 Protease and Inhibition of Assembly and Budding of Virus-like Particles. Virology. doi:10.1006/viro.1993.1174

Karn, J., Martin Stoltzfus, C., 2012. Transcriptional and Posttranscriptional Regulation of HIV-1 Gene Expression. Cold Spring Harb. Perspect. Med. 2, a006916. doi:10.1016/B978-012095440-7/50028-7

Katz, R.A., Greger, J.G., Skalka, A.M., 2005. Effects of cell cycle status on early events in retroviral replication. J. Cell. Biochem. 94, 880–889. doi:10.1002/jcb.20358

Katzmann, D.J., Babst, M., Emr, S.D., 2001. Ubiquitin-dependent sorting into the multivesicular body pathway requires the function of a conserved endosomal protein sorting complex, ESCRT-I. Cell 106, 145–155. doi:10.1016/S0092-8674(01)00434-2

Kaye, J.F., Lever, A.M.L., 1998. Nonreciprocal Packaging of Human Immunodeficiency Virus Type 1 and Type 2 RNA : a Possible Role for the p2 Domain of Gag in RNA Encapsidation Nonreciprocal Packaging of Human Immunodeficiency Virus Type 1 and Type 2 RNA : a Possible Role for the p2 Domain. J. Virol. 72, 5877–5885.

Keane, S.C., Heng, X., Lu, K., Kharytonchyk, S., Ramakrishnan, V., Carter, G., Barton, S., Hosic, A., Florwick, A., Santos, J., Bolden, N.C., McCowin, S., Case, D.A., Johnson, B.A., Salemi, M., Telesnitsky, A., Summers, M.F., 2015. Structure of the HIV-1 RNA packaging signal. Science 348, 917–21. doi:10.1126/science.aaa9266

Keele, B.F., Giorgi, E.E., Salazar-Gonzalez, J.F., Decker, J.M., Pham, K.T., Salazar, M.G., Sun, C., Grayson, T., Wang, S., Li, H., Wei, X., Jiang, C., Kirchherr, J.L., Gao, F., Anderson, J.A., Ping, L.-H.,

Swanstrom, R., Tomaras, G.D., Blattner, W.A., Goepfert, P.A., Kilby, J.M., Saag, M.S., Delwart, E.L., Busch, M.P., Cohen, M.S., Montefiori, D.C., Haynes, B.F., Gaschen, B., Athreya, G.S., Lee, H.Y., Wood, N., Seoighe, C., Perelson, A.S., Bhattacharya, T., Korber, B.T., Hahn, B.H., Shaw, G.M., 2008. Identification and characterization of transmitted and early founder virus envelopes in primary HIV-1 infection. Proc. Natl. Acad. Sci. 105, 7552–7557. doi:10.1073/pnas.0802203105

Keele, B.F., Jones, J.H., Terio, K.A., Estes, J.D., Rudicell, R.S., Wilson, M.L., Li, Y., Learn, G.H., Beasley, T.M., Schumacher-Stankey, J., Wroblewski, E., Mosser, A., Raphael, J., Kamenya, S., Lonsdorf, E. V, Travis, D. a, Mlengeya, T., Kinsel, M.J., Else, J.G., Silvestri, G., Goodall, J., Sharp, P.M., Shaw, G.M., Pusey, A.E., Hahn, B.H., 2009. Increased mortality and AIDS-like immunopathology in wild chimpanzees infected with SIVcpz. Nature 460, 515–519. doi:10.1038/nature08200

Kenyon, J.C., Prestwood, L.J., Le Grice, S.F.J., Lever, A.M.L., 2013. In-gel probing of individual RNA conformers within a mixed population reveals a dimerization structural switch in the HIV-1 leader. Nucleic Acids Res. 41, e174. doi:10.1093/nar/gkt690

Kenyon, J.C., Prestwood, L.J., Lever, A.M.L., 2015. A novel combined RNA-protein interaction analysis distinguishes HIV-1 Gag protein binding sites from structural change in the viral RNA leader. Sci. Rep. 5, 14369. doi:10.1038/srep14369

Kenyon, J.C., Tanner, S.J., Legiewicz, M., Phillip, P.S., Rizvi, T.A., Le Grice, S.F.J., Lever, A.M.L., 2011. SHAPE analysis of the FIV Leader RNA reveals a structural switch potentially controlling viral packaging and genome dimerization. Nucleic Acids Res. 39, 6692–6704. doi:10.1093/nar/gkr252

Kharytonchyk, S., Monti, S., Smaldino, P.J., Van, V., Bolden, N.C., Brown, J.D., Russo, E., Swanson, C., Shuey, A., Telesnitsky, A., Summers, M.F., 2016. Transcriptional start site heterogeneity modulates the structure and function of the HIV-1 genome. Proc. Natl. Acad. Sci. 201616627. doi:10.1073/pnas.1616627113

Kim, H.-J., Lee, K., O'Rear, J.J., 1994. A Short Sequence Upstream of the 5' Major Splice Site Is Important for Encapsidation of HIV-1 Genomic RNA. Virology 198, 336–340.

Kim, J., Sitaraman, S., Hierro, A., Beach, B.M., Odorizzi, G., Hurley, J.H., 2005. Structural basis for endosomal targeting by the Bro1 domain. Dev. Cell 8, 937–947. doi:10.1016/j.devcel.2005.04.001

Klatzmann, D., Barré-Sinoussi, F., Nugeyre, M.T., Dauguet, C., Vilmer, E., Griscelli, C., Brun-Vezinet, F., Rouzioux, C., Gluckman, J.C., Chermann, J.-C., Montagnier, L., 1984. Selective Tropism of Lymphadenopathy Associated Virus (LAV) for Helper-Inducer T Lymphocytes. Science (80-.). 225, 59–63.

Klaver, B., Berkhout, B., 1994. Comparison of 5' and 3' long terminal repeat promoter function in human immunodeficiency virus. J. Virol. 68, 3830–40.

Kohl, N.E., Emini, E.A., Schleif, W.A., Davis, L.J., Heimbach, J.C., Dixon, R.A., Scolnick, E.M., Sigal, I.S., 1988. Active human immunodeficiency virus protease is required for viral infectivity. Proc. Natl. Acad. Sci. U. S. A. 85, 4686–90. doi:10.1073/pnas.85.13.4686

Kol, N., Shi, Y., Tsvitov, M., Barlam, D., Shneck, R.Z., Kay, M.S., Rousso, I., 2007. A Stiffness Switch in Human Immunodeficiency Virus. Biophys. J. 92, 1777–1783. doi:10.1529/biophysj.106.093914

König, R., Zhou, Y., Elleder, D., Diamond, T.L., Bonamy, G.M.C., Irelan, J.T., Chiang, C., Tu, B.P., De Jesus, P.D., Lilley, C.E., Seidel, S., Opaluch, A.M., Caldwell, J.S., Weitzman, M.D., Kuhen, K.L., Bandyopadhyay, S., Ideker, T., Orth, A.P., Miraglia, L.J., Bushman, F.D., Young, J.A., Chanda, S.K., 2008. Global Analysis of Host-Pathogen Interactions that Regulate Early-Stage HIV-1 Replication. Cell 135, 49–60. doi:10.1016/j.cell.2008.07.032

Korber, B., Muldoon, M., Theiler, J., Gao, F., Gupta, R., Lapedes, A., Hahn, B.H., Wolinsky, S., Bhattacharya, T., 2000. Timing the Ancestor of the HIV-1 Pandemic Strains. Science (80-.). 288, 1789–1796. doi:10.1126/science.288.5472.1789

Ku, P.-I., Bendjennat, M., Ballew, J., Landesman, M.B., Saffarian, S., 2014. ALIX is recruited temporarily into HIV-1 budding sites at the end of gag assembly. PLoS One 9, e96950. doi:10.1371/journal.pone.0096950

Kutluay, S.B., Bieniasz, P.D., 2010. Analysis of the initiating events in HIV-1 particle assembly and genome packaging. PLoS Pathog. 6, e1001200. doi:10.1371/journal.ppat.1001200

Kutluay, S.B., Zang, T., Blanco-Melo, D., Powell, C., Jannain, D., Errando, M., Bieniasz, P.D., 2014. Global changes in the RNA binding specificity of HIV-1 gag regulate virion genesis. Cell 159, 1096–1109. doi:10.1016/j.cell.2014.09.057

L'Hernault, A., Greatorex, J.S., Crowther, R.A., Lever, A.M.L., 2007. Dimerisation of HIV-2 genomic RNA is linked to efficient RNA packaging, normal particle maturation and viral infectivity. Retrovirology 4, 90. doi:10.1186/1742-4690-4-90

L'Hernault, A., Weiss, E.U., Greatorex, J.S., Lever, A.M.L., 2012. HIV-2 genome dimerization is required for the correct processing of Gag: a second-site reversion in matrix can restore both processes in dimerization-impaired mutant viruses. J. Virol. 86, 5867–76. doi:10.1128/JVI.00124-12

Langelier, C., von Schwedler, U.K., Fisher, R.D., De Domenico, I., White, P.L., Hill, C.P., Kaplan, J., Ward, D., Sundquist, W.I., 2006. Human ESCRT-II complex and its role in human

immunodeficiency virus type 1 release. J. Virol. 80, 9465–9480. doi:10.1128/JVI.01049-06

Lata, S., Schoehn, G., Jain, A., Pires, R., Piehler, J., Göttlinger, H.G., Weissenhorn, W., 2008. Helical Structures of ESCRT-III Are Disassembled by VPS4. Science (80-.). 321, 1354–1357.

Laughrea, M., Jetté, L., Mak, J., Kleiman, L., Liang, C., Wainberg, M. a, 1997. Mutations in the kissing-loop hairpin of human immunodeficiency virus type 1 reduce viral infectivity as well as genomic RNA packaging and dimerization. J. Virol. 71, 3397–3406.

Laughrea, M., Shen, N., Jetté, L., Wainberg, M.A., 1999. Variant effects of non-native kissing-loop hairpin palindromes on HIV replication and HIV RNA dimerization: Role of stem-loop B in HIV replication and HIV RNA dimerization. Biochemistry 38, 226–234.

Le Grice, S.F.J., 2015. Targeting the HIV RNA Genome: High-Hanging Fruit Only Needs a Longer Ladder, in: The Future of HIV-1 Therapeutics. pp. 147–169.

Lee, K.E., Ambrose, Z., Martin, T.D., Oztop, I., Mulky, A., Julias, J.G., Vandegraaff, N., Baumann, J.G., Wang, R., Yuen, W., Takemura, T., Shelton, K., Taniuchi, I., Li, Y., Sodroski, J., Littman, D.R., Coffin, J.M., Hughes, S.H., Unutmaz, D., Engelman, A., KewalRamani, V.N., 2010. Flexible Use of Nuclear Import Pathways by HIV-1. Cell Host Microbe 7, 221–233. doi:10.1016/j.chom.2010.02.007

Lee, S.-K., Potempa, M., Swanstrom, R., 2012. The choreography of HIV-1 proteolytic processing and virion assembly. J. Biol. Chem. 287, 40867–74. doi:10.1074/jbc.R112.399444

Lever, A.M.L., Gottlinger, H., Haseltine, W., Sodroski, J., 1989. Identification of a Sequence Required for Efficient Packaging of Human Immunodeficiency Virus Type 1 RNA into Virions. J. Virol. 63, 4085–4087.

Lewinski, M.K., Yamashita, M., Emerman, M., Ciuffi, A., Marshall, H., Crawford, G., Collins, F., Shinn, P., Leipzig, J., Hannenhalli, S., Berry, C.C., Ecker, J.R., Bushman, F.D., 2006. Retroviral DNA integration: Viral and cellular determinants of target-site selection. PLoS Pathog. 2, e60. doi:10.1371/journal.ppat.0020060

Li, F., Goila-Gaur, R., Salzwedel, K., Kilgore, N.R., Reddick, M., Matallana, C., Castillo, A., Zoumplis, D., Martin, D.E., Orenstein, J.M., Allaway, G.P., Freed, E.O., Wild, C.T., 2003. PA-457: a potent HIV inhibitor that disrupts core condensation by targeting a late step in Gag processing. Proc. Natl. Acad. Sci. U. S. A. 100, 13555–60. doi:10.1073/pnas.2234683100

Liang, C., Rong, L., Cherry, E., Kleiman, L., Laughrea, M., Wainberg, M.A., 1999a. Deletion Mutagenesis within the Dimerization Initiation Site of Human Immunodeficiency Virus Type 1 Results in Delayed Processing of the p2 Peptide from Precursor Proteins. J. Virol. 73, 6147–

6151.

Liang, C., Rong, L., Laughrea, M., Kleiman, L., Wainberg, M.A., 1998. Compensatory point mutations in the human immunodeficiency virus type 1 Gag region that are distal from deletion mutations in the dimerization initiation site can restore viral replication. J. Virol. 72, 6629–6636.

Liang, C., Rong, L., Quan, Y., Laughrea, M., Kleiman, L., Wainberg, M.A., 1999b. Mutations within four distinct gag proteins are required to restore replication of human immunodeficiency virus type 1 after deletion mutagenesis within the dimerization initiation site. J. Virol. 73, 7014–7020.

Liang, C., Rong, L., Russell, R.S., Wainberg, M.A., 2000. Deletion mutagenesis downstream of the 5' long terminal repeat of human immunodeficiency virus type 1 is compensated for by point mutations in both the U5 region and gag gene. J. Virol. 74, 6251–61. doi:10.1128/JVI.74.14.6251-6261.2000.Updated

Llano, M., Saenz, D.T., Meehan, A., Wongthida, P., Peretz, M., Walker, W.H., Teo, W., Poeschla, E.M., 2006. An essential role for LEDGF/p75 in HIV integration. Science (80-.). 314, 461–464. doi:10.1126/science.1132319

Lu, K., Heng, X., Garyu, L., Monti, S., Garcia, E.L., Kharytonchyk, S., Dorjsuren, B., Kulandaivel, G., Jones, S., Hiremath, A., Divakaruni, S.S., LaCotti, C., Barton, S., Tummillo, D., Hosic, A., Edme, K., Albrecht, S., Telesnitsky, A., Summers, M.F., 2011. NMR detection of structures in the HIV-1 5'-leader RNA that regulate genome packaging. Science (80-.). 334, 242–5. doi:10.1126/science.1210460

Luban, J., Bossolt, K.L., Franke, E.K., Kalpana, G. V, Goff, S.P., 1993. Human immunodeficiency virus type 1 Gag protein binds to cyclophilins A and B. Cell 73, 1067–1078. doi:10.1016/0092-8674(93)90637-6

Maddon, P.J., Dalgleish, A.G., McDougal, J.S., Clapham, P.R., Weiss, R.A., Axel, R., 1986. The T4 gene encodes the AIDS virus receptor and is expressed in the immune system and the brain. Cell 47, 333–348. doi:10.1016/0092-8674(86)90590-8

Mailler, E., Bernacchi, S., Marquet, R., Paillart, J.-C., Vivet-Boudou, V., Smyth, R.P., 2016. The life-cycle of the HIV-1 gag–RNA complex. Viruses 8, 248. doi:10.3390/v8090248

Mak, J., Kleiman, L., 1997. Primer tRNAs for reverse transcription. J. Virol. 71, 8087–95.

Mamede, J.I., Cianci, G.C., Anderson, M.R., Hope, T.J., 2017. Early cytoplasmic uncoating is associated with infectivity of HIV-1. doi:10.1073/pnas.1706245114

Manel, N., Hogstad, B., Wang, Y., Levy, D.E., Unutmaz, D., Littman, D.R., 2010. A cryptic sensor for HIV-1 activates antiviral innate immunity in dendritic cells. Nature 467, 214–217.

doi:10.1038/nature09337

Marquet, R., Baudin, F., Gabus, C., Darlix, J.-L., Mougel, M., Ehresmann, C., Ehresmann, B., 1991. Dimerisation of human immunodeficiency virus (type 1) RNA: stimulation by cations and possible mechanism. Nucleic Acids Res. 19, 2349–2357.

Marquet, R., Paillart, J.-C., Skripkin, E., Ehresmann, C., Ehresmann, B., 1994. Dimerization of human immunodeficiency virus type 1 RNA involves sequences located upstream of the splice donor site. Nucleic Acids Res. 22, 145–151. doi:10.1093/nar/22.2.145

Martin-Serrano, J., Yarovoy, A., Perez-Caballero, D., Bieniasz, P.D., 2003. Divergent retroviral late-budding domains recruit vacuolar protein sorting factors by using alternative adaptor proteins. Proc. Natl. Acad. Sci. 100, 12414–12419.

Martin-Serrano, J., Zang, T., Bieniasz, P.D., 2001. HIV-1 and Ebola virus encode small peptide motifs that recruit Tsg101 to sites of particle assembly to facilitate egress. Nat. Med. 7, 1313–1319. doi:10.1038/nm1201-1313

Matreyek, K.A., Engelman, A., 2011. The Requirement for Nucleoporin NUP153 during Human Immunodeficiency Virus Type 1 Infection Is Determined by the Viral Capsid. J. Virol. 85, 7818–7827. doi:10.1128/JVI.00325-11

Mattei, S., Schur, F.K.M., Briggs, J.A.G., 2016. Retrovirus maturation — an extraordinary structural transformation. Curr. Opin. Virol. 18, 27–35. doi:10.1016/j.coviro.2016.02.008

McBride, M.S., Panganiban, A.T., 1997. Position dependence of functional hairpins important for human immunodeficiency virus type 1 RNA encapsidation in vivo. J. Virol. 71, 2050–8.

McBride, M.S., Panganiban, A.T., 1996. The human immunodeficiency virus type 1 encapsidation site is a multipartite RNA element composed of functional hairpin structures. J. Virol. 70, 2963–2973.

McCallister, S., Lalezari, J., Richmond, G., Thompson, M., Harrigan, R., Martin, D., Salzwedel, K., Allaway, G., 2008. HIV-1 Gag polymorphisms determine treatment response to bevirimat (PA-457), in: Antivir Ther. p. A10.

McDonald, J.H., n.d. T-test for two samples [WWW Document]. Handb. Biol. Stat. URL http://www.biostathandbook.com/twosamplettest.html

McDougal, J.S., Kennedy, M.S., Sligh, J.M., Cort, S.P., Mawle, A., Nocholson, J.K.A., 1986. Binding of HTLV-III/IAV to T4+ T Cells by a Complex of the 110K Viral Protein and the T4 Molecule. Science (80-.). 231, 382–385.

Melikyan, G.B., 2008. Common principles and intermediates of viral protein-mediated fusion: the HIV-1 paradigm. Retrovirology 5. doi:10.1186/1742-4690-5-111

Meng, B., Ip, N.C.Y., Prestwood, L.J., Abbink, T.E.M., Lever, A.M.L., 2015. Evidence that the endosomal sorting complex required for transport-II (ESCRT-II) is required for efficient human immunodeficiency virus-1 (HIV-1) production. Retrovirology 12, 72. doi:10.1186/s12977-015-0197-x

Meng, B., Lever, A.M.L., 2013. Wrapping up the bad news: HIV assembly and release. Retrovirology 10. doi:10.1186/1742-4690-10-5

Miyazaki, Y., Garcia, E.L., King, S.R., Iyalla, K., Loeliger, K., Starck, P., Syed, S., Telesnitsky, A., Summers, M.F., 2010. An RNA Structural Switch Regulates Diploid Genome Packaging by Moloney Murine Leukemia Virus. J. Mol. Biol. 396, 141–152. doi:10.1016/j.jmb.2009.11.033

Moore, M.D., Fu, W., Nikolaitchik, O., Chen, J., Ptak, R.G., Hu, W.-S., 2007. Dimer Initiation Signal of Human Immunodeficiency Virus Type 1 : Its Role in Partner Selection during RNA Copackaging and Its Effects on Recombination. J. Virol. 81, 4002–4011. doi:10.1128/JVI.02589-06

Morgan, D., Mahe, C., Mayanja, B., Martin Okongo, J., Lubega, R., Whitworth, J.A.G., 2002. HIV-1 infection in rural Africa: is there a difference in median time to AIDS and survival compared with that in industrialized countries? AIDS 16, 597–603.

Morita, E., Sandrin, V., Chung, H.-Y., Morham, S.G., Gygi, S.P., Rodesch, C.K., Sundquist, W.I., 2007. Human ESCRT and ALIX proteins interact with proteins of the midbody and function in cytokinesis. EMBO J. 26, 4215–27. doi:10.1038/sj.emboj.7601850

Morita, E., Sandrin, V., McCullough, J., Katsuyama, A., Baci Hamilton, I., Sundquist, W.I., 2011. ESCRT-III protein requirements for HIV-1 budding. Cell Host Microbe 9, 235–242. doi:10.1016/j.chom.2011.02.004

Mougel, M., Tounekti, N., Darlix, J.-L., Paoletti, J., Ehresmann, B., Ehresmann, C., 1993. Conformational analysis of the 5' leader and the gag initiation site of Mo-MuLV RNA and allosteric transitions induced by dimerization. Nucleic Acids Res. 21, 4677–84.

Murakami, T., Ablan, S., Freed, E.O., Tanaka, Y., 2004. Regulation of human immunodeficiency virus type 1 Env-mediated membrane fusion by viral protease activity. J. Virol. 78, 1026–31. doi:10.1128/JVI.78.2.1026

Muriaux, D., Girard, P.M., Bonnet-Mathonière, B., Paoletti, J., 1995. Dimerization of HIV-1 LAI RNA at low ionic strength. J. Biol. Chem. doi:10.1074/jbc.270.14.8209

Muriaux, D., Mirro, J., Harvin, D., Rein, A., 2001. RNA is a structural element in retrovirus particles.

Proc. Natl. Acad. Sci. U. S. A. 98, 5246–51. doi:10.1073/pnas.091000398

Muriaux, D., Rocquigny, H.D., Roques, B.-P., Paoletti, J., 1996. NCp7 Activates HIV-1 Lai RNA Dimerization by Converting a Transient Loop-Loop Complex into a Stable Dimer NCp7 Activates HIV-1 Lai RNA Dimerization by Converting a Transient Loop-Loop Complex into a Stable Dimer *. J Biol Chem. 271, 33686–33692. doi:10.1074/jbc.271.52.33686

Nabel, G., Baltimore, D., 1987. An inducible transcription factor activates expression of human immunodeficiency virus in T cells. Nature 326, 711–3. doi:10.1038/326711a0

Neville, M., Stutz, F., Lee, L., Davis, L.I., Rosbash, M., 1997. The importin-beta family member Crm1p bridges the interaction between Rev and the nuclear pore complex during nuclear export. Curr. Biol. 7, 767–775. doi:10.1016/S0960-9822(06)00335-6

Nikolaitchik, O.A., Dilley, K.A., Fu, W., Gorelick, R.J., Tai, S.-H.S., Soheilian, F., Ptak, R.G., Nagashima, K., Pathak, V.K., Hu, W.-S., 2013. Dimeric RNA recognition regulates HIV-1 genome packaging. PLoS Pathog. 9, e1003249. doi:10.1371/journal.ppat.1003249

Nowicka-Sans, B., Protack, T., Lin, Z., Li, Z., Zhang, S., Sun, Y., Samanta, H., Terry, B., Liu, Z., Chen, Y., Sin, N., Sit, S.-Y., Swidorski, J.J., Chen, J., Venables, B.L., Healy, M., Meanwell, N.A., Cockett, M., Hanumegowda, U., Regueiro-Ren, A., Krystal, M., Dicker, I.B., 2016. BMS-955176: Identification and Characterization of a Second-Generation HIV-1 Maturation Inhibitor with Improved Potency, Anti-viral Spectrum and Gag Polymorphic Coverage. Antimicrob. Agents Chemother. AAC.02560-15. doi:10.1128/AAC.02560-15

Ogawa, M., Takemoto, Y., Sumi, S., Inoue, D., Kishimoto, N., Takamune, N., Shoji, S., Suzu, S., Misumi, S., 2015. ATP generation in a host cell in early-phase infection is increased by upregulation of cytochrome c oxidase activity via the p2 peptide from human immunodeficiency virus type 1 Gag. Retrovirology 12, 97. doi:10.1186/s12977-015-0224-y

Ohagen, A., Gabuzda, D., 2000. Role of Vif in Stability of the Human Immunodeficiency Virus Type 1 Core. J. Virol. 74, 11055–11066. doi:10.1128/jvi.74.23.11055-11066.2000

Onafuwa-Nuga, A.A., Telesnitsky, A., King, S.R., 2006. 7SL RNA, but not the 54-kd signal recognition particle protein, is an abundant component of both infectious HIV-1 and minimal virus-like particles. RNA 12, 542–6. doi:10.1261/rna.2306306

Ott, D.E., Coren, L. V, Shatzer, T., 2009. The nucleocapsid region of human immunodeficiency virus type 1 Gag assists in the coordination of assembly and Gag processing: role for RNA-Gag binding in the early stages of assembly. J. Virol. 83, 7718–27. doi:10.1128/JVI.00099-09

Paillart, J.-C., Berthoux, L., Ottmann, M., Darlix, J.-L., Marquet, R., Ehresmann, B., Ehresmann, C.,

1996. A dual role of the putative RNA dimerization initiation site of human immunodeficiency virus type 1 in genomic RNA packaging and proviral DNA synthesis. J. Virol. 70, 8348–8354.

Paillart, J.-C., Marquet, R., Skripkin, E., Ehresmann, B., Ehresmann, C., 1994. Mutational analysis of the bipartite dimer linkage structure of human immunodeficiency virus type 1 genomic RNA. J. Biol. Chem. 269, 27486–27493.

Paillart, J.-C., Shehu-Xhilaga, M., Marquet, R., Mak, J., 2004. Dimerization of retroviral RNA genomes: an inseparable pair. Nat. Rev. Microbiol. 2, 461–472. doi:10.1038/nrmicro903

Pang, H.-B., Hevroni, L., Kol, N., Eckert, D.M., Tsvitov, M., Kay, M.S., Rousso, I., 2013. Virion stiffness regulates immature HIV-1 entry. Retrovirology 10, 4. doi:10.1186/1742-4690-10-4

Parent, L.J., 2011. New insights into the nuclear localization of retroviral gag proteins. Nucleus 2, 92–97. doi:10.4161/nucl.2.2.15018

Peng, C., Ho, B.K., Chang, T.W., Chang, N.T., 1989. Role of human immunodeficiency virus type 1-specific protease in core protein maturation and viral infectivity. J Virol 63, 2550–2556.

Pettit, S.C., Everitt, L.E., Choudhury, S., Dunn, B.M., Kaplan, A.H., 2004. Initial Cleavage of the Human Immunodeficiency Virus Type 1 GagPol Precursor by Its Activated Protease Occurs by an Intramolecular Mechanism. J. Virol. 78, 8477–8485. doi:10.1128/JVI.78.16.8477-8485.2004

Pettit, S.C., Lindquist, J.N., Kaplan, A.H., Swanstrom, R., 2005. Processing sites in the human immunodeficiency virus type 1 (HIV-1) Gag-Pro-Pol precursor are cleaved by the viral protease at different rates. Retrovirology 2, 66. doi:10.1186/1742-4690-2-66

Piguet, V., Steinman, R.M., 2007. The interaction of HIV with dendritic cells: outcomes and pathways. Trends Immunol. 28, 503–510. doi:10.1016/j.it.2007.07.010

Pincetic, A., Medina, G., Carter, C., Leis, J., 2008. Avian sarcoma virus and human immunodeficiency virus, type 1 use different subsets of ESCRT proteins to facilitate the budding process. J. Biol. Chem. 283, 29822–29830. doi:10.1074/jbc.M804157200

Platt, E.J., Wehrly, K., Kuhmann, S.E., Chesebro, B., Kabat, D., 1998. Effects of CCR5 and CD4 Cell Surface Concentrations on Infections by Macrophagetropic Isolates of Human Immunodeficiency Virus Type 1. J. Virol. 72, 2855–2864.

Poon, D.T.K., Chertova, E.N., Ott, D.E., 2002. Human immunodeficiency virus type 1 preferentially encapsidates genomic RNAs that encode Pr55(Gag): functional linkage between translation and RNA packaging. Virology 293, 368–78. doi:10.1006/viro.2001.1283

Popov, S., Popova, E., Inoue, M., Gottlinger, H.G., 2008. Human Immunodeficiency Virus Type 1 Gag

Engages the Bro1 Domain of ALIX/AIP1 through the Nucleocapsid. J. Virol. 82, 1389–1398. doi:10.1128/JVI.01912-07

Popovic, M., Sarngadharan, M.G., Read, E., Gallo, R.C., 1984. Detection, isolation, and continuous production of cytopathic retroviruses (HTLV-III) from patients with AIDS and pre-AIDS. Science 224, 497–500. doi:10.1126/science.6200935

Prabu-Jeyabalan, M., Nalivaika, E., Schiffer, C.A., 2002. Substrate shape determines specificity of recognition for HIV-1 protease: Analysis of crystal structures of six substrate complexes. Structure 10, 369–381. doi:10.1016/S0969-2126(02)00720-7

Prescher, J., Baumgartel, V., Ivanchenko, S., Torrano, A.A., Brauchle, C., Muller, B., Lamb, D.C., 2015. Super-Resolution Imaging of ESCRT-Proteins at HIV-1 Assembly Sites. PLoS Pathog. 11, e1004677. doi:10.1371/journal.ppat.1004677

Pruss, D., Wolffe, A.P., 1994. Human immunodeficiency virus integrase directs integration to sites of severe DNA distortion within the nucleosome core. Proc. Natl. Acad. Sci. 91, 5913–5917.

Purcell, D.F., Martin, M.A., 1993. Alternative splicing of human immunodeficiency virus type 1 mRNA modulates viral protein expression, replication, and infectivity. J. Virol. 67, 6365–6378.

Rankovic, S., Varadarajan, J., Ramalho, R., Aiken, C., Rousso, I., 2017. Reverse Transcription Mechanically Initiates HIV-1 Capsid Disassembly. J. Virol. 91, e00289-17.

Rasaiyaah, J., Tan, C.P., Fletcher, A.J., Price, A.J., Blondeau, C., Hilditch, L., Jacques, D.A., Selwood, D.L., James, L.C., Noursadeghi, M., Towers, G.J., 2013. HIV-1 evades innate immune recognition through specific cofactor recruitment. Nature 503, 402–405. doi:10.1038/nature12769

Ratner, L., Haseltine, W., Patarca, R., Livak, K.J., Starcich, B., Josephs, S.F., Doran, E.R., Rafalski, J.A., Whitehorn, E.A., Baumeister, K., 1985. Complete nucleotide sequence of the AIDS virus, HTLV-III. Nature 313, 277–84. doi:10.1038/313277a0

Rhee, S., Hunter, E., 1990. A single amino acid substitution within the matrix protein of a type D retrovirus converts its morphogenesis to that of a type C retrovirus. Cell 63, 77–86.

Ristic, N., Chin, M.P.S., 2010. Mutations in matrix and SP1 repair the packaging specificity of a Human Immunodeficiency Virus Type 1 mutant by reducing the association of Gag with spliced viral RNA. Retrovirology 7, 73. doi:10.1186/1742-4690-7-73

Rittner, K., Churcher, M.J., Gait, M.J., Karn, J., 1995. The Human Immunodeficiency Virus Long Terminal Repeat Includes a Specialised Initiator Element which is Required for Tat-responsive Transcription 562–580.

Robb, M.L., Eller, L.A., Kibuuka, H., Rono, K., Maganga, L., Nitayaphan, S., Kroon, E., Sawe, F.K., Sinei, S., Sriplienchan, S., Jagodzinski, L.L., Malia, J., Manak, M., de Souza, M.S., Tovanabutra, S., Sanders-Buell, E., Rolland, M., Dorsey-Spitz, J., Eller, M.A., Milazzo, M., Li, Q., Lewandowski, A., Wu, H., Swann, E., O'Connell, R.J., Peel, S., Dawson, P., Kim, J.H., Michael, N.L., 2016. Prospective Study of Acute HIV-1 Infection in Adults in East Africa and Thailand. N. Engl. J. Med. 374, 2120–2130. doi:10.1056/NEJMoa1508952

Romanelli, M.G., Diani, E., Bergamo, E., Casoli, C., Ciminale, V., Bex, F., Bertazzoni, U., 2013. Highlights on distinctive structural and functional properties of HTLV Tax proteins. Front. Microbiol. 4, 1–14. doi:10.3389/fmicb.2013.00271

Rong, L., Russell, R.S., Hu, J., Guan, Y., Kleiman, L., Liang, C., Wainberg, M.A., 2001. Hydrophobic amino acids in the human immunodeficiency virus type 1 p2 and nucleocapsid proteins can contribute to the rescue of deleted viral RNA packaging signals. J. Virol. 75, 7230–7243. doi:10.1128/JVI.75.16.7230

Rong, L., Russell, R.S., Hu, J., Laughrea, M., Wainberg, M.A., Liang, C., 2003. Deletion of stem-loop 3 is compensated by second-site mutations within the Gag protein of human immunodeficiency virus type 1. Virology 314, 221–8.

Rosen, C.A., Sodroski, J.G., Haseltine, W.A., 1985. The Location of Cis-Acting Regulatory Sequences in the Human T Cell Lymphotropic Virus Type III (HTLV-III / LAV) Long Terminal Repeat. Cell 41, 813–823.

Rosen, C.A., Terwilliger, E., Dayton, A., Sodroski, J.G., Haseltine, W.A., 1988. Intragenice cis-acting art gene repsonsive sequences of the human immunodeficiency virus. Proc. Natl. Acad. Sci. USA 85, 2071–2075.

Roy, B.B., Russell, R.S., Turner, D., Liang, C., 2006. The T12I mutation within the SP1 region of Gag restricts packaging of spliced viral RNA into human immunodeficiency virus type 1 with mutated RNA packaging signals and mutated nucleocapsid sequence. Virology 344, 304–14. doi:10.1016/j.virol.2005.09.011

Russell, R.S., Hu, J., Laughrea, M., Wainberg, M. a, Liang, C., 2002. Deficient dimerization of human immunodeficiency virus type 1 RNA caused by mutations of the u5 RNA sequences. Virology 303, 152–163. doi:10.1006/viro.2002.1592

Russell, R.S., Liang, C., Wainberg, M.A., 2004. Is HIV-1 RNA dimerization a prerequisite for packaging? Yes, no, probably? Retrovirology 1, 23. doi:10.1186/1742-4690-1-23

Russell, R.S., Roldan, A., Detorio, M., Hu, J., Wainberg, M.A., Liang, C., 2003. Effects of a single amino

acid substitution within the p2 region of human immunodeficiency virus type 1 on packaging of spliced viral RNA. J. Virol. 77, 12986–95.

Saad, J.S., Miller, J., Tai, J., Kim, A., Ghanam, R.H., Summers, M.F., 2006. Structural basis for targeting HIV-1 Gag proteins to the plasma membrane for virus assembly. Proc. Natl. Acad. Sci. U. S. A. 103, 11364–9. doi:10.1073/pnas.0602818103

Sakaguchi, K., Zambrano, N., Baldwin, E.T., Shapiro, B. a, Erickson, J.W., Omichinski, J.G., Clore, G.M., Gronenborn, a M., Appella, E., 1993. Identification of a binding site for the human immunodeficiency virus type 1 nucleocapsid protein. Proc. Natl. Acad. Sci. U. S. A. 90, 5219–5223.

Sakuragi, J.-I., Ueda, S., Iwamoto, A., Shioda, T., 2003. Possible Role of Dimerization in Human Immunodeficiency Virus Type 1 Genome RNA Packaging. J. Virol. 77, 4060–4069. doi:10.1128/JVI.77.7.4060-4069.2003

Sakuragi, J., Iwamoto, A., Shioda, T., 2002. Dissociation of Genome Dimerization from Packaging Functions and Virion Maturation of Human Immunodeficiency Virus Type 1 Dissociation of Genome Dimerization from Packaging Functions and Virion Maturation of Human Immunodeficiency Virus Type 1. J. Virol. 76, 959–967. doi:10.1128/JVI.76.3.959

Sakuragi, J., Shioda, T., Panganiban, A.T., 2001. Duplication of the primary encapsidation and dimer linkage region of human immunodeficiency virus type 1 RNA results in the appearance of monomeric RNA in virions. J. Virol. 75, 2557–65. doi:10.1128/JVI.75.6.2557-2565.2001

Sánchez-Luque, F.J., Stich, M., Manrubia, S., Briones, C., Berzal-Herranz, A., 2014. Efficient HIV-1 inhibition by a 16 nt-long RNA aptamer designed by combining in vitro selection and in silico optimisation strategies. Sci. Rep. 4, 1–10. doi:10.1038/srep06242

Schaller, T., Ocwieja, K.E., Rasaiyaah, J., Price, A.J., Brady, T.L., Roth, S.L., Hué, S., Fletcher, A.J., Lee, K., KewalRamani, V.N., Noursadeghi, M., Jenner, R.G., James, L.C., Bushman, F.D., Towers, G.J., 2011. HIV-1 capsid-cyclophilin interactions determine nuclear import pathway, integration targeting and replication efficiency. PLoS Pathog. 7, e1002439. doi:10.1371/journal.ppat.1002439

Schröder, A.R.W., Shinn, P., Chen, H., Berry, C., Ecker, J.R., Bushman, F., 2002. HIV-1 integration in the human genome favors active genes and local hotspots. Cell 110, 521–9. doi:10.1016/s0092-8674(02)00864-4

Schur, F.K.M., Obr, M., Hagen, W.J.H., Wan, W., Jakobi, A.J., Kirkpatrick, J.M., Sachse, C., Kräusslich, H.-G., Briggs, J.A.G., 2016. An atomic model of HIV-1 capsid-SP1 reveals structures regulating

assembly and maturation. Science 353, 506–508. doi:10.1126/science.aaf9620

Sette, P., Dussupt, V., Bouamr, F., 2012. Identification of the HIV-1 NC binding interface in Alix Bro1 reveals a role for RNA. J. Virol. 86, 11608–15. doi:10.1128/JVI.01260-12

Sette, P., O'Connor, S.K., Yerramilli, V.S., Dussupt, V., Nagashima, K., Chutiraka, K., Lingappa, J., Scarlata, S., Bouamr, F., 2016. HIV-1 Nucleocapsid Mimics the Membrane Adaptor Syntenin PDZ to Gain Access to ESCRTs and Promote Virus Budding. Cell Host Microbe 19, 336–48. doi:10.1016/j.chom.2016.02.004

Shah, V.B., Shi, J., Hout, D.R., Oztop, I., Krishnan, L., Ahn, J., Shotwell, M.S., Engelman, A., Aiken, C., 2013. The Host Proteins Transportin SR2/TNPO3 and Cyclophilin A Exert Opposing Effects on HIV-1 Uncoating. J. Virol. 87, 422–432. doi:10.1128/JVI.07177-11

Sharp, P.M., Hahn, B.H., 2011. Origins of HIV and the AIDS Pandemic. Cold Spring Harb. Perspect. Med. 1. doi:10.1101/cshperspect.a006841

Shaw, G.M., Hunter, E., 2012. HIV Transmission. Cold Spring Harb. Perspect. Med. 2.

Shehu-Xhilaga, M., Crowe, S.M., 2001. Maintenance of the Gag / Gag-Pol Ratio Is Important for Human Immunodeficiency Virus Type 1 RNA Dimerization and Viral Infectivity. J. Virol. 75, 1834–1841. doi:10.1128/JVI.75.4.1834

Shen, N., Jetté, L., Liang, C., Wainberg, M.A., Laughrea, M., 2000. Impact of human immunodeficiency virus type 1 RNA dimerization on viral infectivity and of stem-loop B on RNA dimerization and reverse transcription and dissociation of dimerization from packaging. J. Virol. 74, 5729–35.

Shen, N., Jetté, L., Wainberg, M.A., Laughrea, M., 2001. Role of stem B, loop B, and nucleotides next to the primer binding site and the kissing-loop domain in human immunodeficiency virus type 1 replication and genomic-RNA dimerization. J. Virol. 75, 10543–10549. doi:10.1128/JVI.75.21.10543

Short, J.M., Fernandez, J.M., Sorge, J.A., Huse, W.D., 1988. λ ZAP: a bacteriophage λ expression vector with in vivo excision properties. Nucleic Acids Res. 16, 7583–7600.

Siliciano, R.F., Greene, W.C., 2011. HIV latency. Cold Spring Harb. Perspect. Med. 1, 1–20. doi:10.1101/cshperspect.a007096

Skripkin, E., Paillart, J.-C., Marquet, R., Ehresmann, B., Ehresmann, C., 1994. Identification of the primary site of the human immunodeficiency virus type 1 RNA dimerization in vitro. Proc. Natl. Acad. Sci. U. S. A. 91, 4945–4949. doi:10.1073/pnas.91.11.4945

Sleiman, D., Goldschmidt, V., Barraud, P., Marquet, R., Paillart, J.C., Tisné, C., 2012. Initiation of HIV-1

reverse transcription and functional role of nucleocapsid-mediated tRNA/viral genome interactions. Virus Res. 169, 324–339. doi:10.1016/j.virusres.2012.06.006

Sodroski, J., Goh, W.C., Rosen, C., Dayton, A., Terwilliger, E., Haseltine, W., 1986. A second post-transcriptional trans-activator gene required for HTLV-III replication. Nature 321, 412–417. doi:10.1038/321412a0

Sodroski, J., Patarca, R., Rosen, C., Wong-Staal, F., Haseltine, W., 1985. Location of the trans-activating region on the genome of human T-cell lymphotropic virus type III. Science (80-.). 229, 74–77. doi:10.1126/science.2990041

Sodroski, J.G., Rosen, C.A., Haseltine, W.A., 1984. Trans-acting transcriptional activation of the long terminal repeat of human T lymphotropic viruses in infected cells. Science (80-.). 225, 381–385.

Sommerfelt, M.A., Rhee, S.S., Hunter, E., 1992. Importance of p12 protein in Mason-Pfizer monkey virus assembly and infectivity. J. Virol. 66, 7005–11.

Song, R., Kafaie, J., Yang, L., Laughrea, M., 2007. HIV-1 viral RNA is selected in the form of monomers that dimerize in a three-step protease-dependent process; the DIS of stem-loop 1 initiates viral RNA dimerization. J. Mol. Biol. 371, 1084–98. doi:10.1016/j.jmb.2007.06.010

Stephenson, J.D., Li, H., Kenyon, J.C., Symmons, M., Klenerman, D., Lever, A.M.L., 2013. Three-dimensional RNA structure of the major HIV-1 packaging signal region. Structure 21, 951–62. doi:10.1016/j.str.2013.04.008

Strack, B., Calistri, A., Craig, S., Popova, E., Göttlinger, H.G., 2003. AIP1/ALIX is a binding partner for HIV-1 p6 and EIAV p9 functioning in virus budding. Cell 114, 689–699. doi:10.1016/S0092-8674(03)00653-6

Sundquist, W.I., Kräusslich, H.-G., 2012. HIV-1 Assembly, Budding, and Maturation. Cold Spring Harb. Perspect. Med. 2.

Takahashi, K., Baba, S., Chattopadhyay, P., Koyanagi, Y., Yamamoto, N., Takaku, H., Kawai, G., 2000a. Structural requirement for the two-step dimerization of human immunodeficiency virus type 1 genome. RNA 6, 96–102. doi:10.1017/S1355838200991635

Takahashi, K., Baba, S., Hayashi, Y., Koyanagi, Y., Yamamoto, N., Takaku, H., Kawai, G., 2000b. NMR Analysis of Intra- and Inter-Molecular Stems in the Dimerization Initiation Site of the HIV-1 Genome. J. Biochem. 127, 681–686.

Takebe, Y., Seiki, M., Fujisawa, J., Hoy, P., Yokota, K., Arai, K., Yoshida, M., Arai, N., 1988. SR alpha promoter: an efficient and versatile mammalian cDNA expression system composed of the simian virus 40 early promoter and the R-U5 segment of human T-cell leukemia virus type 1

long terminal repeat. Mol. Cell. Biol. 8, 466–472. doi:10.1128/MCB.8.1.466

Taylor, B.S., Sobieszczyk, M.E., McCutchan, F.E., Hammer, S.M., 2008. The Challenge of HIV-1 Subtype Diversity. N. Engl. J. Med. 358, 1590–1602. doi:10.1056/NEJMra0706737

Taylor, R.G., Walker, D.C., Mclnnes, R.R., 1993. E.coli host strains significantly affect the quality of small scale plasmid DNA preparations used for sequencing. Nucleic Acids Res. 21, 1677–1678.

Tersmette, M., Gruters, R.A., de Wolf, F., de Goede, R.E., Lange, J.M., Schellekens, P.T., Goudsmit, J., Huisman, H.G., Miedema, F., 1989. Evidence for a role of virulent human immunodeficiency virus (HIV) variants in the pathogenesis of acquired immunodeficiency syndrome: studies on sequential HIV isolates. J. Virol. 63, 2118–25.

Ueno, T., Tokunaga, K., Sawa, H., Maeda, M., Chiba, J., Kojima, A., Hasegawa, H., Shoya, Y., Sata, T., Kurata, T., Takahashi, H., 2004. Nucleolin and the packaging signal, psi, promote the budding of human immunodeficiency virus type-1 (HIV-1). Microbiol. Immunol. 48, 111–8.

UNAIDS, 2016. UNAIDS Fact Sheet November 2016.

Usami, Y., Popov, S., Göttlinger, H.G., 2007. Potent rescue of human immunodeficiency virus type 1 late domain mutants by ALIX/AIP1 depends on its CHMP4 binding site. J. Virol. 81, 6614–6622. doi:10.1128/JVI.00314-07

Valsamakis, A., Zeichner, S., Carswell, S., Alwine, J.C., 1991. The human immunodeficiency virus type 1 polyadenylylation signal: a 3’ long terminal repeat element upstream of the AAUAAA necessary for efficient polyadenylylation. Proc. Natl. Acad. Sci. U. S. A. 88, 2108–12.

VerPlank, L., Bouamr, F., LaGrassa, T.J., Agresta, B., Kikonyogo, A., Leis, J., Carter, C.A., 2001. Tsg101, a homologue of ubiquitin-conjugating (E2) enzymes, binds the L domain in HIV type 1 Pr55(Gag). Proc. Natl. Acad. Sci. U. S. A. 98, 7724–9. doi:10.1073/pnas.131059198

von Schwedler, U.K., Stray, K.M., Garrus, J.E., Sundquist, W.I., 2003. Functional surfaces of the human immunodeficiency virus type 1 capsid protein. J. Virol. 77, 5439–50. doi:10.1128/JVI.77.9.5439

Von Schwedler, U.K., Stuchell, M., Müller, B., Ward, D.M., Chung, H.Y., Morita, E., Wang, H.E., Davis, T., He, G.P., Cimbora, D.M., Scott, A., Kräusslich, H.G., Kaplan, J., Morham, S.G., Sundquist, W.I., 2003. The protein network of HIV budding. Cell 114, 701–713. doi:10.1016/S0092-8674(03)00714-1

Votteler, J., Sundquist, W.I., 2013. Virus budding and the ESCRT pathway. Cell Host Microbe 14. doi:10.1016/j.chom.2013.08.012.Virus

Wagner, J.M., Zadrozny, K.K., Chrustowicz, J., Purdy, M.D., Yeager, M., Ganser-Pornillos, B.K.,

Pornillos, O., 2016. Crystal structure of an HIV assembly and maturation switch. Elife 5, 1–18. doi:10.7554/eLife.17063

Waki, K., Durell, S.R., Soheilian, F., Nagashima, K., Butler, S.L., Freed, E.O., 2012. Structural and Functional Insights into the HIV-1 Maturation Inhibitor Binding Pocket. PLoS Pathog. 8. doi:10.1371/journal.ppat.1002997

Wang, S.-W., Aldovini, A., 2002. RNA Incorporation Is Critical for Retroviral Particle Integrity after Cell Membrane Assembly of Gag Complexes. J. Virol. 76, 11853–11865. doi:10.1128/JVI.76.23.11853-11865.2002

Wang, T., Tian, C., Zhang, W., Luo, K., Sarkis, P.T.N., Yu, L., Liu, B., Yu, Y., Yu, X.-F., 2007. 7SL RNA Mediates Virion Packaging of the Antiviral Cytidine Deaminase APOBEC3G. J. Virol. 81, 13112–13124. doi:10.1128/JVI.00892-07

Wei, P., Garber, M.E., Fang, S.M., Fischer, W.H., Jones, K.A., 1998. A novel CDK9-associated C-type cyclin interacts directly with HIV-1 Tat and mediates its high-affinity, loop-specific binding to TAR RNA. Cell 92, 451–462. doi:10.1016/S0092-8674(00)80939-3

Wei, X., Decker, J.M., Liu, H., Zhang, Z., Arani, R.B., Kilby, J.M., Saag, M.S., Shaw, G.M., Kappes, J.C., Wu, X., 2002. Emergence of Resistant Human Immunodeficiency Virus Type 1 in Patients Receiving Fusion Inhibitor (T-20) Monotherapy Emergence of Resistant Human Immunodeficiency Virus Type 1 in Patients Receiving Fusion Inhibitor (T-20) Monotherapy. Antimicrob. Agents Chemother. 46, 1896–1905. doi:10.1128/AAC.46.6.1896

Weiss, C.D., Levy, J.A., White, J.M., 1990. Oligomeric Organization of gpl20 on Infectious Human Immunodeficiency Virus Type 1 Particles. J. Virol. 64, 5674–5677.

Weiss, R.A., 1996. Retrovirus classification and cell interactions. J. Antimicrob. Chemother. 1–11.

Wiegers, K., Rutter, G., Kottler, H., Tessmer, U., Hohenberg, H., Kräusslich, H.-G., 1998. Sequential steps in human immunodeficiency virus particle maturation revealed by alterations of individual Gag polyprotein cleavage sites. J. Virol. 72, 2846–54.

Wilkinson, K.A., Gorelick, R.J., Vasa, S.M., Guex, N., Rein, A., Mathews, D.H., Giddings, M.C., Weeks, K.M., 2008. High-throughput SHAPE analysis reveals structures in HIV-1 genomic RNA strongly conserved across distinct biological states. PLoS Biol. 6, e96. doi:10.1371/journal.pbio.0060096

Willey, R.L., Martin, M.A., Peden, K.W., 1994. Increase in soluble CD4 binding to and CD4-induced dissociation of gp120 from virions correlates with infectivity of human immunodeficiency virus type 1. J. Virol. 68, 1029–39.

Wills, J.W., Craven, R.C., 1991. Form, function, and use of retroviral Gag proteins. Aids 5, 639–654.

doi:10.1097/00002030-199106000-00002

Wlodawer, A., Gustchina, A., 2000. Structural and biochemical studies of retroviral proteases. Biochim. Biophys. Acta - Protein Struct. Mol. Enzymol. 1477, 16–34. doi:10.1016/S0167-4838(99)00267-8

Worobey, M., Telfer, P., Souquière, S., Hunter, M., Coleman, C.A., Metzger, M.J., Reed, P., Makuwa, M., Hearn, G., Honarvar, S., Roques, P., Apetrei, C., Kazanji, M., Marx, P.A., 2010. Island Biogeography Reveals the Deep History of SIV. Science (80-.). 329, 1487.

Wright, E.R., Schooler, J.B., Ding, H.J., Kieffer, C., Fillmore, C., Sundquist, W.I., Jensen, G.J., 2007. Electron cryotomography of immature HIV-1 virions reveals the structure of the CA and SP1 Gag shells. EMBO J. 26, 2218–2226. doi:10.1038/sj.emboj.7601664

Wyma, D.J., Jiang, J., Shi, J., Zhou, J., Lineberger, E., Miller, M.D., Aiken, C., Lineberger, J.E., 2004. Coupling of Human Immunodeficiency Virus Type 1 Fusion to Virion Maturation : a Novel Role of the gp41 Cytoplasmic Tail Coupling of Human Immunodeficiency Virus Type 1 Fusion to Virion Maturation : a Novel Role of the gp41 Cytoplasmic Tail. J. Virol. 78, 3429–3435. doi:10.1128/JVI.78.7.3429